أ

أثر حجم العينة وطول الاختبار ونوع الفقرة في دقة تقدير معالم الفقرة والقدرة ودالة معلومات الاختبار

إعداد

نبيل موسى جبران جبران

المشرف

الأستاذ الدكتور عبد الحافظ قاسم الشايب

قدمت هذه الأطروحة استكمالاً لمتطلبات الحصول على درجة الدكتوراه في

علم النفس التربوي / القياس والتقويم

كلية الدراسات العليا

الجامعة الأردنية

آب (أغسطس)/ 2017 م

د.

نموذج ترخيص

أنا الطالب : نبيل موسى جبران جبران أمنح الجامعة الأردنية و/ أو من تفوضه ترخيصاً غير حصري دون مقابل بنشر و / أو استعمال و / أو استغلال و / أو ترجمة و / أو تصوير و / أو إعادة إنتاج بأي طريقة كانت سواء ورقية و / أو إلكترونية أو غير ذلك رسالة الماجستير / الدكتوراه المقدمة من قبلي وعنوانها.

"أثر حجم العينة وطول الاختبار ونوع الفقرة في دقة تقدير معالم الفقرة والقدرة ودالة معلومات الاختبار"

وذلك لغايات البحث العلمي و / أو التبادل مع المؤسسات التعليمية والجامعات و / أو لأي غاية أخرى تراها الجامعة الأردنية مناسبة، وأمنح الجامعة الحق بالترخيص للغير بجميع أو بعض ما رخصته لها.

اسم الطالب: نبيل موسى جبران جبران

التوقيع:

التاريخ: 2017/10/2

ب

قرار لجنة المناقشة

نوقشت هذه الأطروحة (أثر حجم العينة و طول الاختبار ونوع الفقرة في دقة تقدير معالم الفقرة والقدرة ودالة معلومات الاختبار) وأجيزت بتاريخ 3 / 8 / 2017

أعضاء لجنة المناقشة	التوقيع
الدكتور / عبد الحافظ قاسم الشايب، مشرفاً	....
استاذ – القياس والتقويم	
(جامعة آل البيت)	
الدكتور/ أحمد يوسف قواسمه، عضواً خارجياً	....
استاذ – القياس و القويم	
(جامعة اليرموك)	
الدكتورة / فريال محمد أبو عواد، عضواً	....
استاذ مشارك – القياس والتقويم	
(الجامعة الأردنية)	
الدكتور / شفيق فلاح علاونه	....
استاذ – تعلم ونمو	
(الجامعة الأردنية)	

الإهـــــــداء

إلى روح المربي الفاضل ... إلى روح والدي العزيز رحمه الله.

إلى والدتي الغالية ... أمد الله في عمرها.

إلى شريكة حياتي ... إلى زوجتي العزيزة.

إلى أبنائي وبناتي (معاذ، أحمد، جنى، صهيب، جمان) حفظهم الله.

إلى كل طالب علم وباحث يسعى بصدق لخدمة وإسعاد البشرية بعلمه.

إلى الأقارب والأصدقاء والزملاء الأفاضل.

أهدي لكم جميعاً ثمرة عملي وخلاصة جهدي ... مع التحية والتقدير...

نبيل موسى جبران جبران

شكـــر وتقديـــــر

أتقدم بالشكر الجزيل والعرفان بالجميل إلى كل من ساعد وساهم في إنجاز هذا العمل.

فالشكر الكبير للأستاذ الدكتورعبد الحافظ قاسم الشايب الذي أشرف ووجه ونصح ومنحني من وقته وعلمه وخبرته الكثير.

والشكر موصول لأعضاء لجنة المناقشة الأفاضل الذين منحوني من وقتهم وجهدهم في مناقشة هذه الرسالة.

ولا يفوتني أن أشكر معلمي ومعلمات ومديري ومديرات المدارس المتعاونة والمشرفين التربويين والمحكمين الأفاضل.

واللهَ أسأل أن يوفق الجميع لما فيه الخير ..

نبيل موسى جبران جبران

فهرس المحتويات

قائمة الجداول

قائمة الأشكال

قائمة الملاحق

قائمة المختصرات الواردة في الدراسة

الاختصار	المعنى بالانجليزية	المعــــــــــــــــــــــنى
IRT	Item Response Theory	نظرية الاستجابة للفقرة
3PLM	Three–Parameter Logestic Model	النموذج اللوجستي ثلاثي المعلمة
2PLM	Two– Parameter Logistic Model	النموذج اللوجستي ثنائي المعلمة
1PLM	One– Parameter Logisttic Model	النموذج اللوجستي احادي المعلمة
MLE	Maximum Liklihood Estimation	طريقة الارجحية العظمى في التقدير
MAP	Maximum APosteriori	أسلوب القيمة العظمى للتوزيع البعدي في التقدير
SEE	Standard Error of Estimate	الخطأ المعياري في التقدير
	Relative Efficiency	الكفاءة النسبية
EAP	Expected A Posteriori	أسلوب توقع التوزيع البعدي في التقدير
JML	Joint Maximum Likelihood Procedure	أسلوب الأرجحية العظمى المشتركة في التقدير
CML	Conditional Maximum Likelihood	أسلوب الأرجحية العظمى الشرطية في التقدير
MML	Marginal Maximum Liklihood	طريقة الأرجحية العظمى الهامشية في التقدير

أثر حجم العينة وطول الاختبار ونوع الفقرة في دقة تقدير معالم الفقرة والقدرة ودالة معلومات الاختبار

إعـــداد

نبيل موسى جبران جبران

المشرف:

الأستاذ الدكتور عبدالحافظ قاسم الشايب

الملخص

هدفت هذه الدراسة إلى الكشف عن أثر كلٍ من حجم العينة وطول الاختبار وشكل فقرات الاختبار وتفاعلاتها على دقة تقدير معالم الفقرة (الصعوبة، التمييز)، وقدرة الأفراد، ودالة معلومات الاختبار. وتكونت عينة الدراسة من (1015) طالبًا وطالبة من طلبة الصف الثامن الأساسي الملتحقين بمدارس تابعة لقصبة عمان أثناء الفصل الدراسي الأول من العام الدراسي 2016/2015.

أعدّ الباحث اختبار تحصيلي في محتوى مادة الرياضيات للفصل الدراسي الأول؛ تكوّن بصورته الأولية من (96) فقرة من نوع الاختيار من متعدّد ذو الأربعة بدائل. تمّ عرض الاختبار على مجموعة من المحكّمين من ذوي الاختصاص. وبهدف التحقّق من الخصائص السيكومترية للاختبار، وفاعلية فقراته، تمّ تطبيق الاختبار على عينة أوليّة مكوّنة من (163) طالبًا وطالبة من خارج عيّنة الدراسة، حيث تمّ الاحتفاظ بـ (75) فقرة في الاختبار بصورته النهائية في ضوء نتائج التحليل. تمّ تقسيم الاختبار إلى اختبارين؛ اختبار طويل تكوّن من (50) فقرة، وثانٍ قصير تكوّن من (25) فقرة، بعد ذلك تمّ اعادة صياغة فقرات كل من الاختبارين الطويل والقصير الى فقرات من

نوع الصواب والخطأ ليتشكل لدى الباحث أربعة أنواع من الاختبار هي: قصير اختيار من متعدّد، طويل اختيار من متعدّد، قصير صواب وخطأ، طويل صواب وخطأ. تمّ تقسيم العيّنة الكليّة عشوائيًا باستخدام الحاسوب إلى ثلاثة عيّنات جزئية ذات أحجام مختلفة؛ صغيرة، ومتوسطة، وكبيرة حيث تكوّنت كل منها من (250، 500، 1015) مفحوصًا على الترتيب.

تم توظيف النموذج ثنائي المعلمة في سياق نظرية الاستجابة للفقرة لتقدير معلمتي الصعوبة والتمييز للفقرات، ومعلمة القدرة للأفراد، ودالة معلومات الاختبار، ثمّ تمّ استخراج قيمة الخطأ المعياري لهذه التقديرات كمؤشر على دقة التقدير. بعد ذلك تمّ استخدام أسلوب تحليل التباين العاملي (2x3x2) ذو القياسات المتكرّرة أحادي بين الأفراد وثنائي داخل الأفراد.

أشارت نتائج الدراسة إلى أن تقدير معلمة الصعوبة يكون أكثر دقة عند استخدام الاختبار الطويل من نوع الاختيار من متعدّد في عينة كبيرة الحجم، وبينما تبيّن أن تقدير معلمة التمييز يكون أكثر دقة عند استخدام الاختبار الطويل من نوع الاختيار من متعدّد أيضاً لكن في العيّنة صغيرة الحجم. أما بالنسبة لتقدير معلمة القدرة فكانت أكثر دقة عند استخدام الاختبار الطويل من نوع الصواب والخطأ في العيّنة صغيرة الحجم أيضاً، وكان تقدير دالة معلومات الاختبار أكثر دقة عند استخدام الاختبار الطويل من نوع الصواب والخطأ في عيّنة صغيرة الحجم.

<u>الكلمات مفتاحية</u>: حجم العينة، طول الاختبار، نوع الفقرات ، دقة التقدير، معالم الفقرة، القدرة، دالة معلومات الاختبار.

الفصل الأول

مشكلة الدراسة وأهميتها

مقدمة:

تسعى العلوم جميعها بشقيها العلوم الطبيعية والعلوم الإنسانية إلى تطوير أساليب وأدوات قياس تتمتع بمستوىً عالٍ من الدقة والموضوعية، وذلك لقياس الظواهر المتعلقة بها، وبالتالي الوصول إلى دقة وموضوعية في النتائج من أجل تحقيق أهداف العلم من فهم الظواهر وتفسيرها وبالتالي ضبط هذه الظواهر والتحكم فيها، مما جعل عملية القياس وأدواته موضع اهتمام المختصين والعاملين في هذه العلوم.

بالنسبة للعلوم الطبيعية تطورت المقاييس الفيزيائية بشكل كبير في الآونة الأخيرة، وأصبحت الأخطاء التي تصاحب عملية القياس الفيزيائية تكاد لا تذكر مع وجود أدوات قياس متناهية في الدقة تستخدم لقياس الأطوال والأوزان والحجوم والكثافة وغيرها، لكن في المجال الإنساني ما زال الخطأ في القياس يصاحب عملية قياس الكثير من الظواهر الإنسانية مثل قياس الذكاء والتحصيل والاتجاهات والميول والقيم وغيرها من هذه المفاهيم التي لا يتم قياسها بشكل مباشر وإنما قياس ما يدل على وجودها. ومما لا شك فيه أن القياس التربوي والنفسي أكثر صعوبة من القياس الطبيعي، وتأتي هذه الصعوبة من طبيعة الظواهر التربوية والنفسية (الإنسانية) المتداخلة والمتشابكة وتأثرها بالكثير من المتغيرات، فضلاً عن الآراء والأحكام الفردية والذاتية التي تؤثر في عملية قياس الظاهرة التربوية الأمر الذي يتطلب بذل المزيد من الجهد البحثي والضبط الموضوعي واستخدام أدوات قياس تختلف عن تلك المستخدمة في العلوم الطبيعية.

تُعدّ الاختبارات أدوات قياس يتم إعدادها وفق طريقة منظمة مهما كان نوع الاختبار أو الغرض منه، والاختبارات التحصيلية سواء كانت مقننة يتم إعدادها من قبل فريق من المختصين في المناهج والقياس النفسي والتربوي بحيث تقيس مجالاً واسعاً نسبياً من محتوى معين وتتمتع بدرجة عالية من الثبات؛ أو اختبارات يقوم المعلم أو الباحث بإعدادها حسب خطوات واضحة ومدروسة؛ هي من الأدوات الرئيسة الهامة التي يتم اللجوء إليها في المؤسسات التربوية لإثارة الدافعية لدى لطلاب وتقييم ما حققوه من نتاجات تعلم، وتستخدم الاختبارات التحصيلية عموماً لتحقيق أهداف عديدة منها تشخيص الصعوبات وتحديد نقاط الضعف وبالتالي تصميم برامج علاجية للأفراد؛ وتستخدم لأغراض التصنيف كتحديد مسار الطلاب الأكاديمي بما يناسب قدراتهم ومهاراتهم؛ كما تستخدم من قبل المدارس والمؤسسات المهنية لمنح الشهادات لممارسة مهنة معينة؛ كذلك تلجأ المؤسسات التربوية للاختبارات التحصيلية لتقييم مدى النجاح الذي حققته؛ وتستخدم أيضاً لاختيار شخص لوظيفة ما من بين مجموعة من المتقدمين (عودة، 2004).

مما سبق تبرز أهمية الاختبارات واستخداماتها المرهونة بدقتها وصدقها وثباتها وموضوعية النتائج التي تنبثق عنها؛ وهناك أنماط مختلفة من اختبارات التحصيل مثل الاختبارات الشفوية واختبارات المقال، والاختبارات الموضوعية بمختلف أنواعها ولكل منها خصائصها ومميزاتها، والمواقف التي يفضل أن تستخدم فيها (عودة، 2004).

قام علماء القياس المعاصر بجهود بحثية مبتكرة منذ الخمسينات من القرن الماضي؛ لتطوير نظرية سيكومترية يمكن باستخدامها التغلب على كثير من المشكلات التي واجهت النظرية الكلاسيكية في القياس (Classical Test Theory (CTT، ونقص فاعلية هذه النظرية في كثير من التطبيقات العملية المستخدمة، مثل: إنشاء بنوك الأسئلة، والمهام الأدائية، وتصميم أنظمة

اختبارات الموائمة المحوسبة (الاختبارات التكيفية)، وإجراء التكافؤ الأفقي والرأسي للاختبارات، والكشف عن تحيز الفقرات وسلوكها المتمايز، وقياس النمو، وتشخيص صعوبات التعلم (علام، 2005).

ونتيجةً لهذه الجهود جاءت نظرية سيكومترية معاصرة أطلق عليها نظرية الاستجابة للفقرة أو نظرية السمات الكامنة (Item Response Theory (IRT ، وقد تطورت هذه النظرية وما انبثق عنها من نماذج متعددة تطوراُ كبيراُ منذ سبعينيات القرن الماضي وحتى وقتنا الحاضر، وقد استندت هذه التطورات إلى التقنيات المتقدمة، وبرمجيات الحاسوب التي أصبحت متوافرة (علام، 2005).

وتقوم نظرية الاستجابة للفقرة على أساس مفاده أنه يمكن التنبؤ بأداء الأفراد أو تفسير أدائهم في اختبار نفسي أو تربوي معين، في ضوء خاصية أو خصائص مميزة لهذا الأداء تسمى السمات Traits؛ إذ تحاول ربط العلاقة بين مستوى سمة معينة لدى الفرد يقيسها اختبار معين، واستجابته لفقرة من فقرات الاختبار، وذلك عن طريق منحنى خصائص الفقرة (Item (ICC Characteristic Curve الذي يوضح العلاقة المنتظمة بين مستويات السمة المقاسة لدى الأفراد وبين احتمالات الإجابة الصحيحة عن الفقرات (الشريفين، 2012).

ويوصف هذا المنحنى بأنه دالة متزايدة وتيرياً Monotonically Increasing؛ أي أنه كلما زادت قدرة الفرد؛ زاد احتمال إجابته الصحيحة عن الفقرة (Hamleton & Rogers,1988). ويعتمد منحنى خصائص الفقرة (ICC) على ثلاث معالم هي: الصعوبة (b)، والتمييز (a)، والتخمين (c). وتأتي قوة نظرية الاستجابة للفقرة من قوة الافتراضات التي ينبغي تحقيقها في البيانات لتؤدي إلى نتائج موثوقة.

عند التدقيق في هذه الافتراضات القوية والتي لا يمكن تحقيقها بسهولة بواسطة بيانات الاختبار نكتشف قوة هذه النظرية وقوة النماذج اللوجستية الأحادية والثنائية والثلاثية التي تضع افتراضات قوية حول مجموعة فقرات الاختبار، وسنأتي على الحديث عن هذه النظرية وافتراضاتها ونماذجها ومعادلاتها وعلاقاتها الرياضية بشيء من التفصيل في الفصل الثاني لاحقاً.

وبالرغم من ذلك كله تبرز مشكلات وقضايا عديدة تواجه نظريات القياس والنماذج المتعلقة بها عمومًا، ومن أهم هذه القضايا التعامل مع أخطاء القياس وأثرها في تقدير قدرات المفحوصين؛ فالنظرية الجيدة هي التي تقدم أسسًا لفهم أثر هذه الأخطاء وكيفية تقليلها، وتقديم فهماً للعلاقة بين المتغيرات المختلفة؛ وبالتالي تقدير القدرة ضمن مدى ثقة معين. هذا وقد اختلفت طرق تناول أخطاء القياس من نظرية إلى أخرى، فقد افترضت بعضها أن الأخطاء تتوزع توزيعاً طبيعياً في نموذج معين، بينما لا يتم افتراض هذا التوزيع للأخطاء في نموذج آخر، ويتم افتراض أن حجم أخطاء القياس ثابتاً عبر تدريج درجات الاختبار في إطار نظرية، بينما تفترض نظرية أخرى أن حجم أخطاء القياس يرتبط بقدرة المفحوصين الحقيقية (Hambleton & Jones, 1993).

من هنا جاءت هذه الدراسة استكمالاً لما بُذل من جهود تناولت دقة تقدير معالم الفقرات، والقدرة، ودالة المعلومات للفقرة والاختبار، والبحث في المتغيرات التي تؤثر في دقة التقدير ومدى تأثير هذه المتغيرات، حيث تناولت هذه الدراسة ثلاثة متغيرات مقترحة وهي حجم عيّنة المفحوصين، وعدد فقرات الاختبار، ونوع أو شكل هذه الفقرات، بينما تناولت دراسات أخرى متغيرات مختلفة غير هذه المتغيرات.

مشكلة الدراسة:

عند الاطلاع على الأدب النظري يتبين وجود العديد من المشكلات التي تواجه الاختبارات التحصيلية في المجال التربوي والنفسي، والتي تقلل من مصداقية وموثوقية هذه الاختبارات ونتائجها، ولعل أبرز هذه المشكلات ما يأتي:

(1) تأثر الدرجة الكلية للمفحوص بخصائص فقرات الاختبار مثل: الصعوبة، والتمييز، وإمكانية التخمين، والثبات؛ فمعامل صعوبة الفقرة مثلاً يعتمد على قدرة المفحوصين الذين تقدموا للاختبار حيث ترتفع قيمته كلما كان المفحوصون ذوي قدرة مرتفعة، وكذلك الحال بالنسبة لمعامل تمييز الفقرة حيث يتأثر معامل تمييز الفقرة بمدى تجانس قدرات المفحوصين؛ أي أن تقدير معالم الفقرة يعتمد بشكل كبير على قدرات المفحوصين الذين يطبق عليهم الاختبار، وهذا يُعدّ من المآخذ التي تؤخذ على النظرية الكلاسيكية في القياس، التي أشار اليها هامبلتون وسوامينثان (Hambelton & Swaminthan, 1985)، واعتبراها من جوانب قصور عديدة تعاني منها النظرية الكلاسيكية في القياس، بالإضافة إلى جوانب قصور أخرى مثل عدم قدرتها على التقدير الدقيق لقدرات الأفراد ذوي القدرات االمرتفعة، والأفراد ذوي القدرات المتدنية وغيرها من جوانب القصور، لذلك تم استخدام مدخل نظرية الاستجابة للفقرة IRT في هذه الدراسة لما تتمتع به من افتراضات قوية مثل أحادية البعد، والاستقلال الموضعي، وعدم تأثر عملية تقدير معالم الفقرة بخصائص عيّنة المفحوصين، وكذلك عدم تأثر عملية تقدير قدرة المفحوصين بخصائص عّينة الفقرات.

(2) بالرغم من تغلب نظرية الاستجابة للفقرة على جوانب القصور التي عانت منها النظرية الكلاسيكية إلا أنها هي أيضاً تواجه مشاكل؛ منها ما يتعلق بدقة تقدير معالم الفقرة وقدرة الأفراد، حيث هنالك عوامل تؤثر في استقرار ودقة التقدير لهذه المعالم مثل حجم العيّنة، وطول الاختبار،

ونوع فقرات الاختبار المُستخدم، وطريقة التقدير وطريقة اختيار العينة، كما أن دالة معلومات الفقرة هي الأخرى تتأثر بشكل أو بآخر بهذه المتغيرات. من هنا برزت فكرة الدراسة الحالية التي تتناول البحث في تأثر دقة تقدير معالم الفقرة، والقدرة، ودالة المعلومات بعوامل حجم عيّنة المفحوصين، وعدد فقرات الاختبار المُستخدم، ونوع أو شكل الفقرات لدى استخدام اختبار تحصيلي في الرياضيات كمصدر للبيانات.

أسئلة الدراسة:

حاولت هذه الدراسة الإجابة عن الأسئلة الآتية:

1- "ما الأثر الرئيس لكل متغيّر من متغيّرات حجم عيّنة المفحوصين (250، 500، 1015) مفحوصًا، وطول الاختبار (طويل، قصير)، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ)، وأثر التفاعلات الثنائية بين هذه المتغيّرات، والتفاعل الثلاثي بينها في دقة تقدير معالم الفقرة ؟

2- "ما الأثر الرئيس لكل متغيّر من متغيّرات حجم عيّنة المفحوصين (250، 500، 1015) مفحوصًا، وطول الاختبار (طويل، قصير)، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ)، وأثر التفاعلات الثنائية بين هذه المتغيّرات، والتفاعل الثلاثي بينها في دقة تقدير معلمة القدرة ؟

3- "ما الأثر الرئيس لكل متغيّر من متغيّرات حجم عيّنة المفحوصين (250، 500، 1015) مفحوصًا، وطول الاختبار (طويل، قصير)، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ)، وأثر التفاعلات الثنائية بين هذه المتغيّرات، والتفاعل الثلاثي بينها في دقة تقدير دالة معلومات الاختبار ؟

أهمية الدراسة:

تبرز أهمية هذه الدراسة مما يأتي:

أولاً– الأهمية النظرية: تتمثل الأهمية النظرية لهذه الدراسة في توظيفها لنظرية الاستجابة للفقرة من خلال استخدام النموذج اللوجستي ثنائي المعلمة في الكشف عن تأثر درجة تأثر دقة تقدير معالم الفقرة، وقدرة الأفراد، ودالة معلومات الاختبار بعوامل حجم عيّنة المفحوصين، وعدد فقرات الاختبار، ونوع فقرات الاختبار لدى استخدام اختبار تحصيلي في الرياضيات، حيث يشير الأدب النظري إلى أن مثل تلك العوامل قد تؤثر في دقة تقدير معالم الفقرات، وقدرة الأفراد، ودالة معلومات الاختبار لدى توظيف نماذج نظرية القياس الحديثة في تقدير هذه المعالم. ويُشير الأدب النظري في إطار نظرية الاستجابة للفقرة إلى أن هناك علاقة واضحة تربط بين الخطأ المعياري في التقدير، ودالة معلومات الفقرة، ودالة معلومات الاختبار؛ حيث يقل الخطأ المعياري في التقدير كلما توفرت معلومات أكبر حول قدرات المفحوصين. لذلك يتم في إطار نظرية الاستجابة للفقرة اختيار الفقرات التي تزودنا بالقدر الكافي من المعلومات للوصول إلى اختبار جيد يتمتع بدرجة جيدة من دقة القياس عند مستويات قدرة مختلفة.

ثانيًا– الأهمية العملية: تناولت العديد من الدراسات طرقًا إحصائية وصفية للكشف عن درجة تأثر دقة تقدير معالم الفقرة، وقدرات الأفراد، ودالة معلومات الاختبار بعوامل حجم عيّنة المفحوصين، وعدد فقرات الاختبار، ونوع فقرات الاختبار. أما في هذه الدراسة، فقد تم الاهتمام بالتصاميم التجريبية للكشف عن أثر حجم عينة المفحوصين، وعدد فقرات الاختبار، ونوع الفقرات في دقة تقدير معالم الفقرات، وقدرة الأفراد، ودالة معلومات الاختبار، حيث تمّ استخدام التصميم التجريبي من نوع ($3\times2\times2$) للكشف عن الآثار الرئيسة، وأثر التفاعلات بين متغيّرات حجم عيّنة

المفحوصين، وعدد فقرات الاختبار، ونوع فقرات الاختبار في دقة تقدير معالم الفقرات، وقدرة الأفراد، ودالة معلومات الاختبار لدى استخدام اختبار تحصيلي في الرياضيات لقياس تحصيل طلبة الصف الثامن الأساسي. وقد اشتمل حجم العينة على ثلاثة مستويات (250، 500، 1015) فردًا، واشتمل متغيّر عدد فقرات الاختبار (طول الاختبار) على مستويين (25، 50) فقرة، كما اشتمل متغيّر نوع فقرات الاختبار على شكلين (فقرات من نوع الاختيار من متعدد، وفقرات من نوع الصواب والخطأ). من جانبٍ آخر، تم في هذه الدراسة استخدام بيانات حقيقية تم جمعها مباشرة من الميدان، ولم يستند الباحث إلى بيانات مولّدة. وبهذا، يمكن القول أن هذه الدراسة تكتسب أهمية عملية من حيث استخدامها لتصميم تجريبي لاختبار الآثار الرئيسة لثلاثة عوامل (حجم العيّنة، وعدد فقرات الاختبار، ونوع فقرات الاختبار) يُعتقد بأنها تؤثر في دقة تقدير معالم الفقرات، وقدرة الأفراد، ودالة معلومات الاختبار، بالإضافة إلى اختبار أثر التفاعلات الممكنة بين العوامل السابقة في دقة تقدير هذه المعالم.

وبشكل عام تكمن أهمية هذه الدراسة في الجوانب الآتية:

أ- محاولة توفير معلومات حول المتغير الأكثر تأثيرًا في دقة تقدير معالم الفقرات، ومعلمة قدرة المفحوصين، ودالة معلومات الاختبار من بين متغيرات حجم العينة، وعدد فقرات الاختبار، ونوع فقرات الاختبار في إطار النموذج اللوجستي ثلاثي المعلمة، وتوظيف مؤشر الخطأ المعياري في التقدير (SEE) في الحكم على المتغيّر الأكثر دقة في التقدير؛ مما يسهم في إفادة المستخدمين للنظرية الحديثة في القياس وتعرّفهم على المتغيّر الأكثر أثرًا في دقة التقدير من بين هذه المتغيرات.

ب- اثراء الدراسات المهتمة بدقة تقدير معالم الفقرات، وقدرة المفحوصين، ودالة معلومات الاختبار، حيث يُتوقع أن يستفيد من نتائج هذه الدراسة القائمون على إعداد المناهج الأردنية في الرياضيات، والقائمون على إعداد وبناء الاختبارات التحصيلية، وتساعدهم في انتقاء اختبارات تتميز درجاتها بالثبات، وتتصف معالم فقراتها بالدقة، وتقدم معلومات أفضل عن قدرة المفحوصين وبالحد الأدنى من الخطأ.

مصطلحات الدراسة:

فيما يأتي تعريفات اجرائية لمصطلحات الدراسة:

معالم الفقرات: يشير مفهوم معالم الفقرات في إطار هذه الدراسة إلى كلٍ من معلمة الصعوبة ومعلمة التمييز.

معلمة صعوبة الفقرة: هي نقطة على متصل القدرة التي يكون احتمال الإجابة على الفقرة عندها إجابة صحيحة يساوي 0.5 في حالة عدم وجود التخمين للوصول الى الاجابة الصحيحة .

معلمة تمييز الفقرة: هي ميل منحنى خصائص الفقرة الذي يقابل النقطة التي تكون عندها معلمة القدرة تساوي معلمة صعوبة الفقرة والتي يكون عنده احتمال الاجابة الصحيحة = 0.5.

معلمة القدرة: وهي مقدار ما يمتلكه الفرد (السمة الكامنة) التي يستطيع من خلالها الإجابة على فقرات الاختبار، والتي يتم قياسها على متصل السمة التي تتراوح قيمتها نظرياً بين (- ∞،∞)، وعملياً بين (-3،3) لوجت.

اللوجت: وحدة قياس كل من معلمة الصعوبة ومعلمة القدرة للمفحوصين.

دالة معلومات الفقرة: هي تعبير عن مقدار ما تزودنا به الفقرة الاختبارية من معلومات عن قدرة المفحوص، وتكون على شكل اقتران رياضي يعبّر عن كمية المعلومات في تمييز الفقرة بين مستويات القدرة المختلفة.

دالة معلومات الاختبار: هي كمية المعلومات التي يتم حسابها من خلال جمع دوال معلومات فقرات الاختبار عند مستوى قدرة معينة (θ)؛ ويتعلق مفهوم دالة المعلومات بتباين تقدير الأرجحية القصوى لمعلمة القدرة، وكذلك لتقدير معالم الفقرة، فإن تباين هذا التقدير يساوي مقلوب دالة المعلومات المناظرة، والخطأ المعياري لهذا التقدير يساوي مقلوب جذر دالة المعلومات (علام، 2005).

تقدير المعالم: هي عملية تحديد قيم للمعالم أو التعبير الكمي عنها من خلال إجابات الأفراد على الاختبار، ويطلق عادة على هذه العملية مصطلح التعيير (Calibration)، ويصاحب عملية التقدير عادة ما يسمى الخطأ المعياري في التقدير، ويتم الاستعانة ببرامج الحاسوب لإجراء عملية التقدير، وفي هذه الدراسة تم الاستعانة ببرنامج (**Bilog MG**).

دقة التقدير: هو تعبير يشير إلى جودة وكفاءة التقدير، واحتمال أن يكون التقدير أقرب ما يمكن من القيمة الحقيقية، وذلك باستخدام مؤشرات مختلفة منها: الخطأ المعياري في التقدير Standard Error of Estimate (SEE) والذي تم اعتماده في هذه الدراسة.

محددات الدراسة:

- اقتصرت الدراسة على عينة من مجتمع طلبة الصف الثامن في المدارس التابعة لمديرية التربية والتعليم لقصبة عمان (عمان الأولى سابقاً) خلال الفصل الدراسي الأول من العام الدراسي 2016/2015.
- اقتصر الاختبار التحصيلي على فقرات موضوعية من نوع الاختيار من أربعة بدائل للفقرة الواحدة، وفقرات من نوع الصواب والخطأ في مادة الرياضيات.
- تحدّدت فاعلية الاختبار التحصيلي الذي تم استخدامه في هذه الدراسة بالقدر الذي توفرت له دلالات إحصائية مقبولة من الصدق، والثبات، وفاعلية الفقرات.
- تمّ في هذه الدراسة تقدير معالم الفقرات، ومعلمة القدرة، ودالة معلومات الاختبار باستخدام النموذج ثنائي المعلمة (Two Parameter Model (2PLM ، وهو أحد نماذج نظرية الاستجابة للفقرة، واستخدام برنامج (Bilog MG) في التقدير.

الفصل الثاني

الإطار النظري والدراسات السابقة ذات الصلة

الإطار النظري:

يُعدّ القياس من القضايا الرئيسة التي تهتم بها العلوم الطبيعية والعلوم السلوكية والإنسانية على حد سواء، وذلك لأن القياس يستند إلى الأساليب الكمية في التعبير الموضوعي الدقيق عن الظواهر المختلفة. وعلى الرغم من أن مشكلات القياس لا تختلف كثيرًا باختلاف العلوم، إلا أن الباحثين في العلوم السلوكية أكثر اهتمامًا بالأسس المنطقية للقياس من نظرائهم في العلوم الطبيعية. وربما يرجع ذلك إلى صعوبة القياس في العلوم السلوكية، حيث تتميَّز الظواهر السلوكية بالتعقيد، وتعدد المتغيرات وتشابكها.

ويختلف قياس السمات والخصائص الإنسانية، مثل التحصيل، والذكاء، والاستعدادات، والميول، والاتجاهات، وسمات الشخصية عن قياس الخصائص الفيزيائية مثل طول سطح الطاولة، أو كثافة سائل، أو ضغط الغاز. والسمات النفسية كامنة ويستدل عليها من خلال السلوك المرتبط بها مما يجعل قياسها يتم بشكل غير مباشر؛ أي من خلال السلوك المرتبط بهذه السمة، وغير مرتبط بسمات أخرى (علام، 2005).

بناءً على ما تقدم، تبرز الأهمية الخاصة لنظريات القياس والنماذج المتعلقة بها في تطبيقات القياس التربوي والنفسي لأنها تقدم إطارًا للتعامل مع مشكلات القياس، والمشكلات الفنية الأخرى المرتبطة بها. ومن أهم هذه القضايا التعامل مع أخطاء القياس، حيث تقدم النظرية الجيدة

والنموذج الجيد أُسساً لفهم أثر أخطاء القياس في تقدير قدرات المفحوصين وكيفية تقليل هذه الأخطاء، وتقدم النظرية الجيدة كذلك فهمًا للعلاقة بين المتغيرات.

وتتفاوت النظريات المختلفة في طرق تناولها للأخطاء، حيث تفترض إحدى النظريات أن الأخطاء تتوزع توزيعًا طبيعيًا في نموذج معين، بينما لا يتم افتراض مثل هذا التوزيع للأخطاء في نموذج آخر، وكذلك يمكن اعتبار حجم أخطاء القياس ثابتًا عبر تدريج درجات الاختبار، بينما يمكن اعتباره في إطار نظرية أخرى مرتبطًا بالقدرة الحقيقية للمفحوص. ومن ناحية أخرى، فإن نظرية القياس الجيدة تقدم إطارًا مرجعيًا لتصميم الاختبارات ومعالجة مشكلات تطبيقها، وقد تحدد النظرية العلاقة الدقيقة بين فقرات الاختبار وقدرات المفحوصين (Hambleton & Jones, 1993)، وفيما يأتي تعريف عام ومختصر بنظرية الاستجابة للفقرة.

نظرية الاستجابة للفقرة Item Response Theory (IRT):

انبثقت هذه النظرية مما يُسمى نماذج السمات الكامنة في القياس Measurement Latent Trait Models (LTM) المرتبطة بنماذج القياس المعاصرة، حيث يرى أصحاب هذه النماذج أن النماذج الكلاسيكية في بناء الاختبارات تناسب قياس المتغيرات الناجمة عن ملاحظات أو أحكام منفردة مثل: تقدير المعلِّم لطالب في سمة معينة، أو قياس الطول أو الوزن، ولكن يصعب تطبيقها في حالة المتغيرات الناجمة عن قياسات مركبة من ملاحظات غير متكافئة، مثل المجموع الموزون لدرجات مجموعة من الفقرات يشتمل عليها اختبار معين (علام، 2005).

ظهرت نظرية الاستجابة للفقرة كنظرية بديلة للنظرية الكلاسيكية، وبيَّن هامبلتون وسوامينثان (Hambleton & Swaminthan, 1985) ثلاث مزايا رئيسة لنظرية السمات الكامنة، كما يطلق عليها، وهي:

1- يتمّ تقدير قدرة المفحوص بشكل مستقل عن عينة الفقرات التي تطبق عليه؛ أي أن تقدير القدرة للأفراد المفحوصين متحرّرة من خصائص الفقرات المستخدمة في تقدير القدرة (item–free) بافتراض وجود عدد كبير من فقرات الاختبار التي تقيس السمة.

2- يتم تقدير معالم الفقرات (الصعوبة، والتمييز، والتخمين) في أداة القياس (الاختبار أو المقياس) بشكل مستقل عن عينة المفحوصين المُستخدمة في تقدير هذه المعالم (sample–free) بافتراض وجود عدد كبير من المفحوصين.

3- تمكّن النظرية من الحصول على إحصائي statistic يقيس درجة الدقة في قياس قدرة كل مفحوص، وهذا الإحصائي يختلف من مفحوص إلى آخر.

وأضاف علام (2005) مزايا أخرى للنظرية، أهمها:

1- توفّر نماذج النظرية الحديثة أساسًا لربط احتمال الحصول على الإجابة الصحيحة لفقرات الاختبار بمستويات القدرة للأفراد.

2- لا تقوم نماذج النظرية الحديثة على افتراضات غير ممكنة.

3- لا تتطلب نماذج (صور) الاختبار التوازي الصارم لتقدير الثبات والذي يصعب تحقيقه عمليًا.

4- يتم تقدير الخطأ في نماذج النظرية عند كل مستوى من مستويات القدرة بعكس النماذج الكلاسيكية التي تفترض خطأ قياس واحد عند كل مستوى من مستويات القدرة.

افتراضات نظرية الاستجابة للفقرة:

يعتمد أي نموذج رياضي على عدد من الافتراضات التي تتعلق بالبيانات التي يستخدمها النموذج والتي تحدد العلاقات بين الأبنية constructs المُلاحظة وغير المُلاحظة التي يصفها النموذج، لذلك تقوم نظرية الاستجابة للفقرة على الافتراضات الآتية: (Hambleton & Swaminthan, 1985):

1- أحادية البُعد Unidimensionality: يُقصد بذلك أن فقرات الاختبار تقيس قدرة أو سمة واحدة فقط، وعند ذلك يتسنى تفسير إجابة المفحوص على أنها تعود لهذه القدرة أو السمة التي تقيسها فقرات الاختبار. ومن الجدير بالذكر أنه من الممكن للاختبار أن يكون أحادي البُعد في مجتمع، بينما يكون متعدد الأبعاد في مجتمع آخر. ومثال على ذلك عندما يكون الاستيعاب القرائي عاملاً مهمًا في حل المسائل الرياضية في مجتمع معيّن يكون جميع أفراده قادرين على فهم المسائل الرياضية، وتكون القدرة الرياضية هي السمة الوحيدة التي تؤثر في الأداء على الاختبار، بينما في مجتمع آخر يتصف بقدرة أقل في الاستيعاب القرائي من المجتمع الأول بحيث يكون أفراده غير قادرين على استيعاب المسائل الرياضية في الاختبار. في هذه الحالة تكون مهارة الاستيعاب القرائي والقدرة الرياضية كلاهما تؤثران في أداء الأفراد، وبالتالي يكون هذا الاختبار متعدد الأبعاد Multidimensional. ويُعد افتراض أحادية البُعد هو أكثر افتراضات نظرية القياس الحديثة (IRT) تعقيداً وتقييداً، ومن الطبيعي أن لا يتحقق هذا الافتراض لوجود عوامل أخرى غير القدرة تؤثر في أداء المفحوص منها الدافعية، قلق الاختبار، السرعة في الأداء، معرفة استخدام ورقة الاختبار،...الخ. لذا فإن المطلوب لتحقيق هذا الافتراض

هو وجود عامل سائد Dominant Factor واحد يؤثر في الأداء على الاختبار، ويتم التحقق من هذا الافتراض بإجراء تحليل عاملي Factor Analysis للفقرات، ومن الخطأ الاعتقاد بأن القدرة المقاسة بالاختبار لا تتغير مع مرور الزمن، بل إنها تتأثر بالتعلم، والنسيان، وعوامل أخرى؛ فهي متغيرة بشكل مستمر. وعليه فإن نظرية السمات الكامنة العامة تفترض وجود عدة سمات كامنة تفسر أداء المفحوصين على مجموعة فقرات الاختبار، وهذه المجموعة من السمات الكامنة تعرَف أبعاد فضاء السمات الكامنة Dimensional Latent Space. ويمكن تحديد موقع كل مفحوص على كل سمة من مجموعة السمات من خلال تحديد موقع المفحوص في فضاء السمات الكامنة، ويكون فضاء السمة الكامنة كاملاً complete إذا كانت كل السمات تؤثر في درجات الاختبار لمجموعة المفحوصين.

2- الاستقلال الموضعي Local Independence: يُقصد بهذا الافتراض أن استجابة المفحوصين من نفس القدرة على أي زوج من فقرات الاختبار تكون مستقلة عن بعضها إحصائيًا، وهذا يعني أن أداء الفرد على فقرة من فقرات الاختبار لا يؤثر على أدائه على أية فقرة أخرى على افتراض أن كل فقرة مستقلة في بنائها، ولا توحي بالإجابة عن أية فقرة أخرى، أي أن الأفراد الذين يملكون القدرة نفسها يجب أن يكون معامل الارتباط بين أدائهم على فقرة وأدائهم على فقرة أخرى مساوياً للصفر. وعندما يتحقق افتراض الاستقلال الموضعي يكون احتمال الحصول على أي نمط معين لدرجات فرد ما مساويًا لحاصل ضرب احتمالات حدوث هذا النمط لكل فقرة من فقرات الاختبار التي أجاب عنها، فمثلاً إذا كانت إجابة مفحوص على خمس فقرات كالآتي (u=1,0,1,1,0) حيث 1: تعني

إجابة صحيحة على الفقرة، 0: تعني إجابة خاطئة، فإن احتمال الحصول على هذا النمط من الاستجابة هو:

$$P_1(1-P_2)(P_3)(P_4)(1-P_5) \quad -$$

حيث: P_1 هي احتمال إجابة المفحوص إجابة صحيحة على الفقرة الأولى، (1−P2) احتمال إجابة المفحوص إجابة خاطئة على الفقرة الثانية. ويرى هامبلتون Hambelton أن افتراض الاستقلال الموضعي مكافئ لفرض أحادية البُعد للسمة الكامنة، فإذا لم تكن الفقرات مستقلة إحصائيا عند مستوى معيّن من القدرة فهذا يعني أن العلامات المتوقعة لبعض المفحوصين على الاختبار ستكون أعلى مقارنة ببعض المفحوصين ممن هم في المستوى نفسه من القدرة. وبالتالي فإن الاختبار يقيس أكثر من سمة واحدة، وهذا انتهاك لافتراض أحادية البُعد. وعندما تزودنا الفقرة بإيحاءات تساعدنا على الإجابة عن فقرة أخرى فإنه يمكن أن يكتشف بعض المفحوصين هذه الإشارة، والبعض الآخر قد لا يكتشفها، وهنا تصبح القدرة على كشف الإشارة هي البُعد الذي يفحص وليست القدرة (Hambelton & Jones, 1993). وتجدر الإشارة هنا إلى أن افتراض الاستقلال الموضعي لا يتضمن أن الفقرات غير مترابطة عبر مجموعة المفحوصين الكلية؛ بل على العكس يجب أن تكون مترابطة لأنها تقيس السمة نفسها، وهذا يشبه إلى حد ما مبادئ التحليل العاملي حيث يفترض أن تكون المتغيرات مترابطة نظرًا لقياسها السمة نفسها، ولكن بعد إجراء التحليل نخرج بعوامل مستقلة، وهذا يعني بأن العلامات غير مترابطة عند مستوى قدرة معين وليست الفقرات (Hambleton & Swaminathan, 1985).

3- منحنى خصائص الفقرة (ICC) Item Characteristic Curve : تفترض نماذج الاستجابة للفقرة أحادية البُعد وجود متصل للسمة المراد قياسها، بحيث يمكن تقدير احتمال إجابة فرد إجابة صحيحة على فقرة اختبارية ما إذا علمنا موقعه على هذا المتصل. ويعبّر عن ذلك رياضيًا بأن الاختبار يُعد بمثابة راسم فوقي Mapping من عينة الأفراد المفحوصين إلى متصل السمة الكامنة، واحتمال الإجابة الصحيحة يكون دالة متزايدة مطردة لموقع الأفراد على متصل السمة، وبذلك يزداد احتمال توصل الفرد إلى الإجابة الصحيحة على الفقرة الاختبارية بزيادة السمة لديه والتي تمثل هنا القدرة. ومنحنى خصائص الفقرة هو دالة رياضية تربط بين احتمال الإجابة الصحيحة على الفقرة بقدرة المفحوص التي يقيسها الاختبار، وتكون على شكل دالة انحدار غير خطية، وتكون هذه الدالة مميزة وخاصة بكل فقرة من فقرات الاختبار يتخذ كل منها شكل منحنى الترجيح اللوغاريتمي الاحتمالي Logistic Curve، ويسمى أحيانًا "المنحنى المميز للفقرة"، هذا إذا كانت الفقرة ثنائية الدرجة Dichotomous Item كما هو الحال في فقرات الاختيار من متعدّد، فإن المنحنى يعبّر عن "انحدار احتمال الإجابة الصحيحة" على "مستوى السمة أو القدرة"، أما إذا كانت الفقرة متعدّدة التدريج Polytomous Item كما في مقاييس الاتجاهات، فإن هذا المنحنى يعبّر عن "انحدار احتمال الاستجابات لكل قسم من أقسام المقياس" على "مستوى السمة". أما في حالة الفقرة ثنائية البُعد؛ أي الفقرة التي تقيس سمتين مختلفتين، فإن المنحنى يصبح سطحاً يسمى السطح المميز للفقرة Surface Characteristic Curve. وبما أن منحنى خصائص الفقرة تراكميًا فإنه يشير بوضوح إلى أن احتمال إجابة الفقرة إجابة صحيحة يزداد بازدياد قدرة المفحوص. وعادة توصف هذه المنحنيات في نماذج الاختبارات التي تقيس سمة واحدة (أحادية البُعد) بدلالة معلمة

الصعوبة، أو معلمتي الصعوبة والتمييز للفقرة، أو معالم الصعوبة والتمييز والتخمين للفقرة، حيث يرجع الاختلاف الرئيس بين نماذج السمات الكامنة المستخدمة إلى اختلاف شكل منحنى خصائص الفقرة. وتقاس القيم التقديرية لكل من قدرة الأفراد، وصعوبة الفقرات بوحدة قياس مشتركة تُسمى "وحدة الترجيح اللوغاريتمي أو اللوجت Logit"؛ وأن احتمال الإجابة الصحيحة على فقرة بواسطة مفحوص ما لا يعتمد على عدد المفحوصين الذين يقعون في نفس مستوى القدرة، وهذه الخاصية التي تشير إلى ثبات منحنيات خصائص الفقرة لمجتمع من المفحوصين يتم معايرة الفقرات لهم هي من أهم مميزات نماذج السمات الكامنة. وفي معظم تطبيقات نظرية الاستجابة للفقرة يأخذ هذا المنحنى شكل الحرف (S) حيث إن زيادة قيمة السمة الكامنة والتي تمثل القدرة (θ) التي يمثلها المحور الافقي يتبعه زيادة في احتمال الإجابة الصحيحة على الفقره $P_i(\theta)$ الممثلة على المحور العمودي، وهو منحنى مطرد تزداد فيه احتمالية الإجابة الصحيحة بازدياد القدرة، ويقترب خط الاستواء السفلي فيه من الصفر عند انعدام التخمين بالإجابة، فيما يقترب خط الاستواء العلوي من الواحد الصحيح. ويكون مستوى صعوبة الفقرة ممثلاً على متصل القدرة الأفقي، ويقابل احتمال الإجابة الصحيحة المساوي للقيمة (0.5) عندما لا تتأثر الإجابة بالتخمين العشوائي، وتمثل تلك النقطة نقطة انقلاب المنحنى. أما القدرة التمييزية أو معلمة التمييز للفقرة فهي تمثل ميل المنحنى عند نقطة الانقلاب، أي بالقرب من المتوسط الذي يقابل النقطة على متصل القدرة التي تمثل الصعوبة. ويدل الميل الحاد steep slope على أن التغير الصغير نسبيًا في قدرة الفرد يجعل احتمال إجابته الصحيحة على الفقرة يتباين تباينًا ملحوظًا (أي أن التمييز مرتفع). أما الميل البسيط فيدل على أن التغيير الكبير نسبيًا في قدرة الفرد لا يغير احتمال الإجابة الصحيحة تغييرًا

ملحوظًا (أي أن التمييز منخفض). لذلك فإن تمييز الفقرة يتناسب تناسبًا طرديًا مع ميل المنحنى المميز لها عند نقطة انقلاب المنحنى. ويتضح مما سبق أن المنحنى المميز للفقرة يعد العمود الفقري لنظرية الاستجابة للفقرة (Hammbleton & Swaminthan, 1985).

4- سرعة الأداء (Speedness): تفترض نماذج نظرية السمات الكامنة أن الاختبارات التي تطابقها هذه النماذج لا تُطبّق تحت شرط السرعة في الأداء على الاختبار؛ بمعنى أن إخفاق الفرد في إجابته على الاختبار يرجع إلى انخفاض قدرته وليس إلى تأثير عامل السرعة على إجابته. وهذا الافتراض نادرًا ما يتم الإشارة إليه كونه متضمناً في افتراض أحادية البعد؛ فعندما تؤثر السرعة على أداء الفرد للاختبار يكون هناك قدرتان على الأقل تؤثران على أدائه هما: سرعة الأداء والسمة المقاسة بواسطة محتوى الاختبار .

نماذج الاستجابة للفقرة:

يذكر ماكدونالد McDonald المشار إليه في (Hambleton & Swaminthan, 1985) نماذج رياضية عامة تتضمن الاعتبارات الآتية:

1- نماذج أحادية البُعد وأخرى متعددة الأبعاد.

2- نماذج خطية وأخرى غير خطية.

3- نماذج استجابات ثنائية ونماذج استجابات غير ثنائية.

تهدف نماذج الاستجابة للفقرة إلى تحديد أو نمذجة العلاقة بين مستوى الفرد في سمة كامنة يقيسها اختبار معين، واستجابته لفقرات الاختبار. لذلك يُطلق على هذه النماذج "الدوال

الاحتمالية للاستجابة للفقرة Item Response Function، حيث تحدد في صيغ رياضية احتمال الاستجابة الصحيحة لفرد ذي مستوى قدرة معين، والصيغة الرياضية العامة لهذه الدالة هي:

$$P_i(\theta) = \frac{e^x}{1+e^x} \dots\dots\dots\dots\dots [1]$$ -

حيث يشير $P_i(\theta)$ إلى احتمال الاستجابة الصحيحة لفرد قدرته (θ) على الفقرة (i)، ويمثل e الأساس اللوغاريتم الطبيعي، وهو مقدار ثابت = 2.718 تقريبًا ويسمى العدد النيبيري، أما x فهي رمز اختياري يعبر عنه بالمعالم المناسبة لأي من هذه النماذج (علام، 2005).

خصائص نماذج الاستجابة للفقرة:

- أداء المفحوص على اختبار معين يمكن التنبؤ به وتفسيره اعتمادًا على واحدة أو أكثر من السمات الكامنة.
- يحدد نموذج الاستجابة للفقرة العلاقة بين أداء المفحوص على اختبار معين والسمة الكامنة التي تكمن وراء هذا الأداء أو السلوك.
- نموذج الاستجابة للفقرة الجيد يزودنا بتقدير أو تنبؤ ذو معنى عن درجات المفحوصين على السمات.
- يتم التنبؤ بالسمة وتقديرها من خلال أداء المفحوص المُلاحظ، ويعبر عن مقدار ما يمتلكه المفحوص من هذه السمة بالقدرة (θ) (Hambleton & Swaminthan, 1985).

أنواع نماذج الاستجابة للفقرة أحادية البعد:

هناك نماذج متعددة في إطار نظرية الاستجابة للفقرة، وذلك لاختلاف الافتراضات المتعلقة بالبيانات الاختبارية، لذلك يمكن تقسيم هذه النماذج إلى قسمين رئيسيين:

- نماذج الاستجابة للفقرة أحادية البعد Unidimensional Models
- نماذج الاستجابة للفقرة متعددة الأبعاد Multidimensional Models

وينقسم كل من النوعين السابقين إلى نماذج سكونية Static Models ونماذج ديناميكية ، Dynamic Models، والذي يهمنا هنا النماذج السكونية (علام، 2005).

النماذج السكونية:

يهتم هذا النوع من النماذج بالقياس في مدة زمنية واحدة، ويهتم أيضًا بتحديد العمليات التي ينطوي عليها الأداء في الاختبارات السيكلوجية والتربوية. وتختلف هذه النماذج في صيغتها الرياضية، كما تختلف في عدد الخصائص أو المعالم المستخدمة في وصف كل فقرة. وتطبق هذه الصيغ في حالة الفقرات ثنائية الاستجابة Dichotomous وكذلك في حالة الفقرات متعددة الاستجابة Polytomous. ونظرًا للاهتمام في هذه الدراسة بالاختبارات التحصيلية المشتملة على فقرات ثنائية الاستجابة، فسيتم التركيز على النماذج المتعلقة بهذا النوع من الفقرات.

النماذج المتعلقة بالفقرة ثنائية الاستجابة Dichotomous Item:

عندما تكون الفقرة ثنائية الاستجابة فإن منحنى خصائص الفقرة يعبر عن انحدار احتمال الإجابة الصحيحة على مستوى السمة أو القدرة كما في فقرات الاختيار من متعدد أو الصواب والخطأ، حيث تتطلب مثل هذه الفقرات إجابة صحيحة أو خاطئة فقط، وتضع قيود وتحد من تنوع صيغ التقويم (Hambleton & Swaminthan, 1985). وهناك نوعين من النماذج التي تتعامل مع هذا النوع من الفقرات: النماذج اللوجستية، والنماذج المعتمدة على المنحنى الطبيعي

(السوي) في بنيتها، لكن الاهتمام سيكون – في هذه الدراسة– منصبًا على النماذج اللوجستية لشيوعها وسعة انتشارها.

النماذج اللوجستية:

تفترض هذه النماذج وجود دالة مميزة وخاصة بكل فقرة يتخذ كل منها شكل منحنى Logistic curve وتقسم حسب عدد معالم الفقرة (الصعوبة، التمييز، التخمين) في كل نموذج إلى:

1– النموذج اللوجستي أحادي المعلمة (One– Parameter Logisttic Model, 1PLM):

ويسمى أيضاً نموذج "راش" (Rash Model). ويُعدّ هذا النموذج أبسط نماذج الاستجابة للفقرة أحادية البعد واشتهر بإسم عالم الرياضيات الدنماركي جورج "راش" بجامعة كوبنهاجن، الذي اطلق اسم " نماذج القياس Measurement Models" على مجموعة من النماذج الإحصائية، والنموذج الذي يعتمد على الدرجة الثنائية هو واحد منها، أي بمعنى أن نموذج "راش" لا يعد نموذجًا واحدًا.

تمثل منحنى خصائص الفقرة في هذا النموذج دالة ترجيح لوغاريتمي أحادية المعلم (أو أحادية البارامتر One – Parameter L ogistic Function، وتكون المتغيرات المستقلة في هذا النموذج قدرة الفرد (θ_j)، وصعوبة الفقرة (b_i)، والمتغير التابع هو احتمال أن الفرد (j) يجيب إجابة صحيحة على الفقرة الاختبارية (i) ويرمز لهذا الاحتمال بالرمز $P_i(\theta)$ ويتم نمذجتها بصيغة رياضية احتمالية، وتُعد دالة ترجيح لوغاريتمي غير خطية كالآتي:

$$P_i(\theta) = \frac{e^{\theta_j - b_i}}{1 + e^{\theta_j - b_i}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots . [2] \quad -$$

وتسمى هذه الصيغة "نموذج الترجيح اللوغاريتمي أحادي المعلمة"، ويفترض هذا النموذج بالإضافة للافتراضات السابقة تساوي القوة التمييزية وتساوي (1) لجميع فقرات الاختبار، أي أن منحنيات خصائص فقرات الاختبار تكون ذات ميل واحد، أما صعوبة الفقرة (b_i) فتمثل النقطة على متصل القدرة التي تقابل احتمال إجابة المفحوص إجابة صحيحة على الفقرة وتساوي (0.5) (Baker, 1985).

يوضح نموذج "راش" ما يأتي (علام، 2001):

- معلمة قدرة الفرد (θ) تقيس ما تقيسه معلمة صعوبة الفقرة (b) ويعبر عنه على نفس المقياس.
- تعتبر نقطة الصفرعلى تدريج المقياس كنقطة مرجع تعود إليها تقديرات كل من قدرة الفرد وصعوبة الفقرة.
- الذي يحكم احتمال الاستجابة الصحيحة أو الخاطئة هو الفرق بين القدرة وصعوبة الفقرة، فإذا كانت $\theta = b$ يكون احتمال الإجابة الصحيحة = 0.5.

هناك العديد من الدراسات في الأدب التربوي التي أجريت بشكل خاص على نموذج ""راش""، وقد أجمعت في معظمها على أن استخدام هذا النموذج لا يتطلب عددًا كبيرًا من المفحوصين، ويمكن استخدامه عند وجود عدد قليل من الفقرات، بحيث يزودنا بدرجة عالية من الدقة في تقدير قدرة الفرد ومعامل صعوبة الفقرة. وأشار الدرابيع (2001) إلى أن استخدام هذا النموذج يزودنا بدقة أعلى في تقدير قدرة الفرد عند استخدام اختبار مكون من (10) فقرات خاصة عندما يكون الأفراد من ذوي القدرات المتدنية، وفي دراسات أخرى يقدم تقديرات أفضل للقدرة

ومعامل الصعوبة عند استخدام حجم عينة قليل يتراوح بين (25−50) فردًا، وعدد قليل من الفقرات (الدرابيع، 2001) .

2− النموذج اللوجستي ثنائي المعلمة Two− Parameter Logistic Model, 2PLM:

اقترح هذا النموذج عالم الإحصاء بيرنبوم (Birnbaum, 1986) مع مجموعة زملائه بجامعة كولومبيا الأمريكية، ويسمى أيضاً نموذج بيرنبوم، ويختلف عن نموذج "راش" السابق بأنه يسمح للفقرات بأن تختلف في كل من الصعوبة والتمييز، إذ يصعب الحصول على فقرات اختبارية متساوية بالقوة التمييزية بين مستويات السمة أو القدرة التي يقيسها الاختبار، وهو الافتراض الذي استند اليه نموذج "راش"، وكذلك فإن الأداء على الفقرات لا يكون من خلال التخمين، أي أن التخمين في هذا النموذج غير موجود، بل تختلف الفقرات في معلمتي الصعوبة والتمييز فقط، لكن يفتقر هذا النموذج إلى بعض الخصائص الإحصائية التي يتميّز بها نموذج "راش"، لذلك فإن عملياته الحسابية أكثر صعوبة (علام، 2005 ؛ Hambleton & Swaminathan, 1985).
ويمثل النموذج الصورة الرياضية الآتية:

$$P_i(\theta) = \frac{e^{\alpha_i(\theta_j - b_i)}}{1+e^{\alpha_i(\theta_j - b_i)}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [3] \quad -$$

وتضمين معلمة التمييز α_i في هذا النموذج يجعل تقدير القدرة معتمدًا على النمط العام لإجابات الأفراد على الفقرة؛ إذ يمكن أن نحصل على تقديرين مختلفين لمستوى السمة من نمطين مختلفين ودرجاتهما الكلية واحدة؛ فتمكّن الفرد من الإجابة الصحيحة على فقرات تمييزها مرتفع، وإخفاقه في الإجابة الصحيحة على فقرات تمييزها منخفض يؤدي إلى تقدير أعلى لمستوى القدرة. لذلك فإن هذا النموذج لا يعيّن أوزاناً متساوية في تقدير مستويات القدرة بعكس نموذج "راش"،

ويترتب على ذلك إمكانية تقاطع منحنيات خصائص الفقرات للاختبار المُعدّ حسب هذا النموذج. وتختلف صعوبة الفقرات في النموذج الثنائي بطريقة ما عن صعوبتها في نموذج "راش"، حيث إن ميل المنحنى عند النقطة المقابلة لصعوبة الفقرة يمثل تمييز تلك الفقرة؛ فكلما زاد ميل المنحنى دل ذلك على زيادة في تمييز الفقرة، علما بأن الميول والمنحدرات المقدرة تتفاوت بشكل كبير بين المنحنيات المميزة لخصائص الفقرات، وتزداد صعوبة الفقرة بنقصان تمييزها (Embretson & Steven Reise,2000؛ Hambelton & Swaminthan, 1985).

3- النموذج اللوجستي ثلاثي المعلمة Three-Parameter Logestic Model, 3PLM:

يُسمى أيضاً نموذج "لورد Lord". ويفترض النموذج أن المفحوصين ذوي القدرة المنخفضة قد يلجؤون للتخمين العشوائي عند الإجابة على فقرات اختبار الاختيار من متعدّد، مما يزيد من احتمال إجابتهم إجابة صحيحة، وهذا ما يحصل في الواقع فعلاً، لذلك أضاف هذا النموذج معلمة أخرى هي معلمة التخمين (c)، لتصبح المعالم في هذا النموذج ثلاثة معالم هي الصعوبة (b)، التمييز(α) والتخمين (c)، والصيغة الرياضية لهذا النموذج هي :

$$P_i(\theta) = C_i + (1 - C_i)\frac{e^{\alpha_i(\theta_j - b_i)}}{1+e^{\alpha_i(\theta_j - b_i)}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots.[4] \quad -$$

وتسمى المعلمة (c) أيضاً بالخط التقاربي الأدنى (Lower Asymptote) لمنحنى خصائص الفقرة، ويمثل احتمال الإجابة الصحيحة للأفراد ذوي القدرة المنخفضة. والهدف من إضافة هذه المعلمة في النموذج هو محاولة مراعاة عدم مطابقة منحنيات خصائص الفقرة عند الطرف الأدنى لمتصل القدرة، حيث يكون المتغير العشوائي أحد المتغيرات التي تؤثر في الأداء على المتغيرات (علام،2005). وعند مطابقة البيانات الاختبارية لهذا النموذج قد يتطلب ذلك أن تكون قيمة معلمة التخمين أصغر من القيمة الناتجة عن احتمال توصل الأفراد من ذوي القدرة

المنخفضة إلى الإجابة الصحيحة على الفقرة عن طريق التخمين العشوائي. وهذا يعتمد على المهارة في كتابة مشتتات (Distractors) فقرات الاختيار من متعدد. لذلك يفضل لورد (Lord) تسمية هذه المعلمة باسم آخر هو "معلمة مستوى شبه الصدفة" "Pseudo-Chance Level Parameter" (Lord & Novick, 1968؛ علام، 2005). وفي الفقرات من نوع الاختيار من متعدد ذي الأربعة بدائل فإن احتمالية الوصول إلى الإجابة الصحيحة بالتخمين العشوائي يساوي 0.25، وغالبًا تختلف التقديرات من النموذج الثلاثي خلال التخمين العشوائي؛ فمثلاً إذا استخدم المفحوصون منهجية الاستبعاد للإجابة غير الصحيحة وغير المنطقية فإن احتمالية الوصول للإجابة الصحيحة بهذه الطريقة هي أعلى من طريقة التخمين العشوائي. ويمكن أن يقودنا النموذج الثلاثي إلى مشاكل في التقدير، ولتجنب بعض المشاكل فإن التخمين المشترك غالبًا ما يقدر للفقرات المتشابهة، ويجب الأخذ بعين الاعتبار أن صعوبة الفقرة لها معانٍ مختلفة في النموذج الثلاثي، إذ إن صعوبة الفقرة التي يكون فيها التخمين (0.25) ترتبط بشكل كبير مع احتمالية أكثر من (0.5) للإجابة الصحيحة (Embretson & Steven Reise, 2000).

النماذج الديناميكية:

تهتم هذه النماذج بقياس التغيّر الذي يحدث في السمات الإنسانية عبر الزمن، حيث يرى البعض أن هذا التغيّر يعد عملية تدريجية، بينما يرى آخرون أنه عملية تحوّل من حالة إلى أخرى. ويكون المتغير الذي يُحدث التغيّر بالاعتماد عليه في الحالة الأولى متصلاً، بينما يكون في الحالة الثانية منفصلاً، غير أن الحالة الأولى هي التي نالت الاهتمام في النماذج الديناميكية. ويشمل النموذج الذي اقترحه بوك (Bock) المشار إليه في (علام، 2005) على عدة مكونات استنادًا إلى النموذج ثنائي المعلمة. وتتطلب عملية التعيير (Calibration) باستخدام هذا النموذج توافر عينة

متكافئة من المفحوصين الذين حصلوا على درجات متفاوتة من التعليم (Instruction)، ويمكن استخدام هذا النموذج في قياس مستويات التحصيل.

تقدير المعالم للنماذج أحادية البُعد:

إن الخطوة الأولى والأساسية في تصميم الاختبارات والمقاييس في سياق نظرية الاستجابة للفقرة هي تقدير المعالم التي يشتمل عليها النموذج الذي تم اختياره، ويعتمد ذلك على توافر أساليب وإجراءات مناسبة لتقدير هذه المعالم، وبرامج محوسبة فعّالة ومناسبة، ويطلق على عملية الحصول على هذه التقديرات مصطلح " التعيير Calibration" (علام، 2005). ويعتمد نجاح عملية التقدير هذه على توفير إجراءات وطرق مناسبة، وعادة تكون معالم الفقرة ومعالم الأفراد مجهولة، وتجدر الإشارة إلى وجود أسلوبين للتقدير هما: أسلوب التقدير بنقطة Point Estimation، وأسلوب التقدير بفترة Interval Estimation. ويتم في الأسلوب الأول استخدام قيمة التقدير بدلاً عن المعلمة المجهولة، لكن لوجود بعض الأخطاء في عملية التقدير وضرورة معرفة مدى انحراف أو اختلاف التقدير عن قيمة المعلمة الحقيقية، يتم استخدام أسلوب التقدير بفترة Interval Estimation (Casella & Berger,1986). وتجدر الإشارة إلى أن هناك تقدير لمعالم الأفراد (القدرة) وتقدير آخر يتعلق بمعالم الفقرات (الصعوبة، والتمييز، والتخمين)، ولكلٍ طرقه وإجراءاته القائمة على افتراضات معينة.

تقدير قدرات الأفراد:

يتم تقدير القدرة أو السمة الكامنة خلف إجابات المفحوصين على مجموعة من الفقرات الاختبارية معالمها معلومة من خلال نمط استجابة المفحوص، ويتم التقدير بإحدى الطرق الآتية:

1- **طريقة الأرجحية العظمى (MLE) Maximum Liklihood Estimation**: ينطلق هذا الأسلوب من إيجاد قيمة القدرة (θ) التي تعطي أرجحية عظمى للحصول على نمط الإجابة عن الفقرات المعطاة، بحيث تأخذ الإجابة الصحيحة الدرجة (1)، وتأخذ الإجابة الخاطئة الدرجة (0)، ويرمز للإجابة ثنائية الدرجة هذه بالرمز"u"، واحتمال الإجابة الصحيحة (P_i)، وتمثّل المعادلة الآتية دالة الأرجحية العظمى:

- $$L(u/\theta) = \prod_{i=1}^{n} P_i^{u}\, Q_i^{1-u} \ldots\ldots\ldots\ldots\ldots\ldots\ldots [5]$$

حيث:

L (u/θ): دالة الأرجحية العظمى

u: نمط إجابة المفحوص ويأخذ القيم {1,0}

θ: قدرة الفرد التي سيتم تقديرها

i: رقم الفقرة

n: عدد الفقرات

ويكون احتمال الفشل في الإجابة الصحيحة Q_i حيث $Q_i = 1 - P_i$

وتحسب دالة الأرجحية العظمى عند أكثر من قيمة أو مستوى للقدرة، وقيمة القدرة (θ) التي تعطي أكبر قيمة لدالة الأرجحية العظمى هي قدرة المفحوص، فمثلاً إذا كان نمط الإجابة لأحد الأفراد المفحوصين عن خمس فقرات هي (1،1،1،0،0) على الترتيب فإن اقتران الأرجحية العظمى يعطى بالعلاقة الآتية :

$$L(u/\theta) = P_1(\theta) P_2(\theta) P_3(\theta) Q_4(\theta) Q_5(\theta)$$

وتجدر الإشارة إلى أن عملية ضرب الاحتمالات في دالة الأرجحية العظمى لا تتم إلا إذا كان هناك استقلالية في الإجابة، أي أن افتراض الاستقلال الموضعي متحقق، ونظرًا لأن الاحتمال

يقع في الفترة [0−1] فإن حاصل الضرب لاحتمال الحصول على أي من أنماط الاستجابات يعطي قيمة صغيرة، ويكون فيها التقريب كبيرًا مما يزيد من الوقوع في أخطاء التقريب، لذلك يتم اللجوء إلى عملية إيجاد اللوغاريتم الطبيعي للأرجحية والذي يحوّل ناتج ضرب الاحتمالات إلى ناتج جمع هذه الاحتمالات، مما يسهل عملية التعامل مع الأرجحية؛ إذ إن اقتران الأرجحية في المعادلة السابقة يصبح اقتران لوغاريتم الأرجحية على الصورة :

- $$L(u/\theta) = \prod_{i=1}^{n} P_i^{u}\, Q_{\mathbf{i}}^{\mathbf{1-u}} \ldots\ldots\ldots\ldots\ldots..[6]$$

يصبح اقتران لوغاريتم الأرجحية على الصورة:

$$lnL(u/\theta) = \sum_{i=1}^{n}[u_{\,i} lnP_i + (1-u) lnQ_i] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [7]$$

وبسبب تحديد قيمة θ التي يكون لاقتران اللوغاريتم الطبيعي للأرجحية عندها قيمة عظمى، ولأن عدد الفقرات التي تمثل النمط المعطى لاستجابات أي مفحوص يكون كبيرًا، لذلك لا بد من استخدام أسلوب أو خوارزمية معينة في تقدير معالم الأفراد ومعالم الفقرات، ومن أهم هذه الخوارزميات وأشهرها: خوارزمية نيوتن – رافسون (Newton – Raphson) التي تعتمد على استخدام المشتقة الأولى $\acute{L}(\theta)$، والمشتقة الثانية $L''(\theta)$ لاقتران لوغاريتم الأرجحية، ويتم تطبيق هذه الخوارزمية من خلال الخطوات الآتية:

- البدء بتقدير أولي للقدرة (θ) بناءً على نمط الإجابة حيث يرمز لهذا التقدير بالرمز (θ_o).
- إيجاد النسبة :

- $$\epsilon_{\mathbf{o}} = \frac{\acute{L}(\boldsymbol{\theta})}{\mathbf{L''(\theta)}} \ldots\ldots\ldots\ldots..[8]$$

والتي تمثل النسبة بين المشتقة الأولى لاقتران الأرجحية وقيمة المشتقة الثانية عند التقدير المعطى للقيمة θ، والمشتقة الأولى والثانية على الترتيب تتمثلان بالصيغ الرياضية الآتية:

- $$\acute{L}(\boldsymbol{\theta}) = \sum_{i=1}^{L}(\alpha_i(u_i - P(\theta)) \ldots\ldots\ldots.[9]$$

$$\mathbf{L}''(\boldsymbol{\theta}) = \sum_{\mathbf{i=1}}^{\mathbf{L}}[\boldsymbol{\alpha}_{\mathbf{i}}^{\mathbf{2}}(\mathbf{1} - \mathbf{P}(\boldsymbol{\theta})(-\mathbf{P}(\boldsymbol{\theta})] \dots\dots\dots [10]$$ -

- إيجاد التقريب الأفضل للقدرة θ من خلال العلاقة : $\theta_1 = \theta_o - \epsilon_o$.
- وضع التقدير θ_1 بديلاً عن التقدير θ_o في الدورة التالية من التقدير.
- تكرار الخطوات من 1−4 حتى يتم الوصول إلى أن يكون الفرق بين قيمة θ التي يتم إيجادها والقيمة التي تسبقها مباشرة أقل من معيار معين يتم تحديده مسبقاً

(Hambleton & Swaminthan, 1985 ؛ Embretson & Steven Reise, 2000؛ التقي، 2009).

خصائص طريقة الأرجحية العظمى لتقدير القدرة:

- تتصف بالقوّة والاتساق، أي كلما زاد عدد الفقرات وزاد عدد المفحوصين كلما اقتربت θ المقدرة من θ الحقيقية.
- يقترب توزيع القدرة المقدرة من التوزيع السوي.
- لا تحتاج هذه الطريقة إلا لنمط إجابة المفحوص كي يتم تقدير قدرته، ولا تحتاج إلى معلومات مسبقة عن توزيع القدرة في مجتمع المفحوصين.
- الفاعلية أو الكفاية: أي أن أدوات تقدير الأرجحية العظمى وبشكل متقارب يكون لها التباين الأصغر، والكفاية بشكل عام تعني أن توزيع العينة لا يعتمد على المعلمة المراد تقديرها.

لكن يوجد محددات لهذه الطريقة من أهمها: أنه عندما يتم تقدير قدرة مفحوص أجاب إجابة كاملة، أو لم يجب على أية فقرة إجابة صحيحة فيتم تقدير قدرته بهذه الطريقة إما +∞ للأول أو−∞

للثاني، وللتغلب على هذا العيب تم اقتراح طرق أخرى للتقدير (Hambleton & Swaminthan, 1985).

ويذكر امبريتسون و ريس (Embretson & Steven Reise, 2000) أن أسلوب الأرجحية العظمى من الأساليب التي تستخدم مفهوم التقاربية في تقدير السمة الكامنة، وبالتالي فإن هذا الأسلوب يحتاج إلى عينة كبيرة من الفقرات. وقد أكد هذا الرأي كل من ثيسن ووينر (Thissen & Wainer) المشار إليهما في عبابنة (2004) حيث أوضحا أنه وبسبب الطرق الرياضية المتبعة في حساب الخطأ المعياري لتقديرات معالم النماذج اللوجستية، وللحصول على تقدير دقيق لمعلمة الصعوبة باستخدام طريقة الأرجحية العظمى لغايات معادلة الاختبارات أو القياس التكيّفي مثلاً فإن حجم العينة اللازم لمدى مقبول من الدقة قد لا يمكن تحقيقه في معظم التطبيقات. ويضيف هامبلتون و سوامينثان (Hambleton & Swaminthan, 1985) أن تقدير الأرجحية العظمى للقدرة تحصل تقاربيًا على أقل تباين يمكن الحصول عليه بحيث يكون التباين الأدنى مساويًا $[I(\theta)]^{-1}$ أي مقلوب دالة المعلومات، وبالتالي لا يمكن أن يكون لأي أسلوب تقدير آخر للقدرة تباين أصغر من ذلك تقريباً. وتجدر الإشارة إلى استخدام المفهوم التقاربي في التقدير ويمكن أن لا تكون صالحة لأي حجم عينة، وأن النتائج المتقاربة تكون صالحة فقط عندما يكون عدد الفقرات (وليس عدد المفحوصين) أكبر، ذلك لأننا معنيون بتقدير θ عندما تكون معالم الفقرة معلومة.

يمكن القول أن الدقة في تقدير المعالم تعتمد على مبدأين أساسيين وهما:

- عدم التحيّز في التقدير، أي أن القيم التقديرية المتوقعة أو قيمة متوسط جميع العينات العشوائية ينبغي أن تساوي قيمة معلمة المجتمع. ويشير التحيّز إلى توقع الفرق بين تقدير المعلمة وقيمتها الفعلية، ولا يعني تحقق هذا المبدأ أن التقدير دقيق تمامًا.
- معدل مربعات الأخطاء للتقدير(Expected Mean Squared Error , EMSE) والذي يتم قياسه بواسطة تباين القيمة المقدّرة، ويُعد أسلوباً لاختيار تقدير جيد للمعالم (Casella & Berger, 1986).

تجدر الاشارة الى أن الباحث في هذه الدراسة قد اعتمد (طريقة الأرجحية العظمى) في تقدير المعالم من خلال برنامج (Bilog MG) نظراً لشيوع استخدامها من قبل الكثير من الباحثين في دراساتهم، ولخصائصها المتعددة وفاعلية هذه الطريقة وكفايتها في عملية التقدير كما تُعد طرق التقدير الأخرى امتداداً لهذه الطريقة والتي تشكل الأساس لطرق التقدير الأخرى، وبغض النظر عن عيوبها التي لا تُذكر أمام خصائصها ومميزاتها.

2- طرق تقدير تعتمد على نظرية بييز Bayes Theorm: تعتمد هذه الطرق على نظرية بييز والتي تربط بين الاحتمالات الشرطية conditional probabilities، والاحتمالات الهامشية marginal probabilities. وتستخدم هذه الطرق عندما لا نستطيع تطبيق طريقة الأرجحية العظمى، والذي يميّز هذا الأسلوب استخدامه لمعلومات سابقة عن العينة تسمى المعلومات الأولية priori information والمتوفرة من خبرات سابقة، ويتم اعتبار كمية ذات تباين معروف من خلال وصف التوزيع الاحتمالي للقدرة، يطلق عليه التوزيع القبلي priori distribution. ويعتمد تحديده على معرفة الباحث وخبرته السابقة عن قدرات الأفراد، وذلك قبل أن يتم جمع البيانات، لذلك أطلق عليه التوزيع القبلي، ويتم مراجعته من خلال المعلومات المتوفرة من العينة، ويطلق على التوزيع

الجديد اسم التوزيع البعدي postreriori distribution. ويتم تقدير قدرات الأفراد بأسلوبين (Embretson & Steven Reise, 2000; Lord, 1986). يكتفي الباحث بذكرهما دون الخوض بالتفاصيل:

(أ) التقدير بأسلوب القيمة العظمى للتوزيع البعدي (Maximum APosteriori, MAP):

وتسمى أيضاً طريقة بييز الشرطية، أو تعظيم الاقتران MAP، وتستخدم نفس الأسلوب والإجراءات المستخدمة في طريقة الأرجحية العظمى حيث يتم التقريب المتتابع لإعطاء تقديرات ثابتة لمعلمة القدرة. إلا أنه يتم استعمال التوزيع القبلي للسمة الكامنة أو القدرة، والذي يكون متوسطه مساويًا للصفر، وانحرافه المعياري مساويًا للواحد الصحيح، وبالعادة يكون التوزيع الطبيعي المعياري، ويتم تعريف التوزيع البعدي من خلال ضرب اقتران الأرجحية باقتران التوزيع القبلي، ويعرف التوزيع البعدي عند التطبيق على أساس اقتران لوغاريتم الأرجحية مضافًا إليه لوغاريتم التوزيع القبلي، تعمل هذه الطريقة حيثما تتوقف طريقة الأرجحية العظمى عن العمل.

(Lord, 1986 ؛Hambleton & Swaminthan, 1985).

(ب) التقدير باستخدام أسلوب توقع التوزيع البعدي (Expected A Posteriori , EAP):

ويطلق على هذه الطريقة أيضاً توقع الاقتران البعدي، أو تقدير بييز Bayesian Estimation، وقد اقترح هذا الأسلوب في التقدير كل من بوك ومسليفي (Bock & Mislevy, 1982) حيث يختلف هذا الأسلوب عن الأسلوبين السابقين من طرق التقدير؛ ففي هذا الأسلوب لا يتم استخدام التقدير الدوراني الذي يعتمد أسلوب نيوتن – رافسون، وإنما يستخدم طريقة مباشرة.

ويتم حساب الخطأ المعياري في التقدير من خلال الجذر التربيعي لمعكوس دالة المعلومات (Mislevy & Bock, 2003; Lord, 1986). وتختلف عملية التقدير في الطرق

الثلاث باختلاف نوعية الفقرات التي تم الإجابة عنها، وبالتالي يمكن ملاحظة أهمية درجة التمييز في تقدير السمة الكامنة. وتوصل الشريفين (2012) في دراسته إلى أن دقة تقديرات معلمة القدرة تزداد في حالة عينة الأفراد ذوي القدرة العالية وعينة الأفراد ذوي القدرة المتدنية عند استخدام طريقة بييز للتوقع (EAP) في حين تزداد الدقة عند مستويات الأفراد ذوي القدرة المتوسطة باستخدام طريقة الأرجحية العظمى (MLE) بغض النظر عن حجم العينة.

تقدير معالم الفقرات:

في تطبيقات (IRT) عادة تكون معالم الفقرات ومستويات القدرة مجهولة ويجب تقديرها من نفس البيانات، ويكون تقدير معالم الفقرات (الصعوبة b، التمييز α، التخمين c) مع عدم معرفة مستويات القدرة يشبه الحصول على انحدار لوجيستي مع عدم معرفة القيم المتنبأ بها (Embretson & Steven Reise, 2000).

وتهدف أساليب تقدير المعالم لتحديد قيمة القدرة (θ) لكل فرد مختبر أو مفحوص، وكذلك قيم معالم الفقرة، من خلال إجابات الأفراد على هذه الفقرات، وهذا أيضاً يماثل تقدير قيم معاملات الانحدار من الاستجابات الملاحظة، غير أن نماذج الاستجابة للفقرة تختلف عن نماذج الانحدار التي تكون عادة خطية (Linear) بينما نماذج الاستجابة غير خطية (Nonlinear)، كما أن المتغيرات المستقلة في تحليل الانحدار يمكن ملاحظتها، بعكس المتغير المستقل في نماذج الاستجابة للفقرة وهو القدرة (θ) الذي لا يكون مُلاحظًا.

ويُلاحظ أنه عندما يصبح عدد الفقرات وعدد المفحوصين كبيرًا فإن تقديرات الأرجحية العظمى لمعالم الفقرات ومعالم المفحوصين تقترب من قيمها الحقيقية (Hambleton & Swaminethan, 1985). ومع وجود العديد من الأساليب والمنهجيات لتيسير تقدير معالم

نماذج الاستجابة للفقرة، توجد محكات ينبغي الاستناد إليها في تقييم مدى جودة إحصائي استدلالي معين، حيث يصاحب عملية تقدير معالم الفقرات والقدرة للأفراد أخطاء في التقدير تعود إلى مصادر مختلفة، لذا يجب توخي الدقة بحيث تكون هذه الأخطاء في حدها الأدنى قدر الإمكان، وقد أورد لورد (Lord,1980) مجموعة من المؤشرات أو المحكات منها:

- عدم التحيز Unbiasedness: وهذا يعني أن القيم التقديرية المتوقعة أو قيمة متوسط جميع العينات العشوائية الممكنة ذات حجم معين، ينبغي أن تساوي قيمة معالم المجتمع؛ بمعنى أن تكون القيم المقدرة للمعالم قريبة من القيم الحقيقية لتلك المعالم، مما يدل على عدم التحيَّز في التقدير، وكلما اقترب مؤشر التحيَّز من الصفر دلَ ذلك على دقة أعلى في التقدير.
- الجذر التربيعي لمتوسط مربعات الخطأ Root Mean Squared Error (RMSE) حيث يتم إيجاد هذا المؤشر أو المحك من خلال إيجاد الجذر التربيعي لمربع الفرق بين القيمة المقدرة والقيمة الحقيقية لكلٍ من معالم الفقرات أو معلمة القدرة، وكلما اقتربت قيمته من الصفر دلَ ذلك على دقة أعلى للتقدير .
- الاتساق Consistency: يعني أن قيمة الإحصائي تقترب تدريجيًا من قيمة معلم المجتمع كلما زاد حجم العينة.
- الفعالية النسبية Relative Efficiency: أي إذا توفر مقياسان إحصائيان غير متحيّزان لتقدير معلم المجتمع فإن أفضلهما هو الأكثر فعالية بالنسبة للآخر، أي يكون الخطأ المعياري لتوزيع معايناته أقل.

- الكفاية Sufficiency: وتعني أن توزيع العينة لا يعتمد على المعلم المراد تقديره، أي أن توزيع أسلوب تقدير معالم المجتمع يشتمل على جميع مفردات العينة المتعلقة بالتقدير، ويكون كافياً لتزويدنا بهذه المعلومات (علام، 2005).

ومن بين أهم أساليب تقديرات الأرجحية العظمى في تقدير معالم الفقرات والقدرة ما يأتي:

(أ) أسلوب الأرجحية العظمى المشتركة Joint Maximum Likelihood Procedure (JML):

تُعد تقديرات (JML) إجراءات تكرار للتقدير مبنية على تقديرات الأفراد ومعالم الفقرات،

ومن مميّزات طريقة الأرجحية العظمى المشتركة في التقدير (JML) أن الخوارزمية سهلة البرمجة. وأن هذه الطريقة تنطبق على العديد من نماذج (IRT) ثنائية وثلاثية المعالم. ومع أن هذه الطريقة تعتمد على أسلوب مباشر في عملية التقدير لا يتناول أياً من الافتراضات لكل من توزيع القدرة أو توزيع معالم الفقرات، إلا أن هناك بعض السلبيات لهذه الطريقة نوجزها بما يأتي (Embretson & Steven Reise, 2000 ؛ Hambleton & Swaminthan, 1985):

- قد يكون لاقتران لوغاريتم الأرجحية أكثر من قيمة عظمى مما قد يعطي تقديرًا خاطئًا لأي من معالم الفقرات إذا كانت القيم التي يتم البدء بها بعيدة عن القيم التي يراد تقديرها.
- عدم وجود قيمة عظمى للوغاريتم الأرجحية في الحالات التي تكون فيها الإجابة عن جميع الفقرات كاملة، أو في الحالة التي تكون فيها الإجابة عن جميع الفقرات معدومة أو خاطئة، وبالتالي لا تستطيع هذه الطريقة إيجاد التقديرات المطلوبة لكل من الفقرات والقدرات، فالعلامات المتطرفة لا تزودنا بمعلومات عن المعالم.

- مع أن ثبات التقدير يحتاج إلى عدد كبير من الأفراد إلا أن زيادة عدد الأفراد تزيد من عدد القدرات المطلوب تقديرها، وبالتالي قد لا تحسّن بالشكل المقبول عملية التقدير؛ فتقديرات معالم الفقرات متحيّزة لكن معامل التصحيح يلغي معظم التحيّز، على أن تقدير معالم الفقرات غير متسق في الاختبارات التي لها طول ثابت.
- عدم الدقة في حساب الخطأ في التقدير لأن التقديرات تعتمد على تقديرات غير متسقة ومتحيّزة.

(ب) أسلوب الأرجحية العظمى الشرطية (Conditional Maximum Likelihood (CML:

تشترط هذه الطريقة شرطًا يُعد كافيًا لتقدير معالم القدرة عند الأفراد، وهو مجموع الفقرات التي أجاب عنها الفرد، وشرطًا آخر يُعد كافياً لتقدير معالم الفقرات، وهو مجموع الأفراد الذين أجابوا عن فقرة معينة إجابة صحيحة. ومن هنا فإن هذه الطريقة تشترط حذف الفقرات التي أجاب عنها جميع الأفراد إجابة صحيحة، أو التي لم يتمكن جميع الأفراد من الإجابة عنها، وكذلك حذف الأفراد الذين أجابوا إجابة صحيحة عن جميع الفقرات، أو الأفراد الذين لم يجيبوا عن أية فقرة من الفقرات. والنموذج الذي يتوافر فيه هذا الشرط هو نموذج "“راش”" أحادي المعلمة، لذا تنحصر هذه الطريقة في هذا النموذج (Hambleton & Swaminthan, 1985).

أما من حيث مميزات وعيوب هذه الطريقة، فقد لخّصها ايمبريتسون و ريس (Embretson & Steven Reise, 2000; Baker,1987) بما يأتي:

مميزات أسلوب الأرجحية العظمى الشرطية:

- لا تتطلب افتراض مسبق لتوزيع القدرة في تقدير درجة الصعوبة؛ فمعالم القدرة المجهولة لا تؤخذ بعين الاعتبار في معادلات تقدير معلم صعوبة الفقرات.

- تقديرات معالم الفقرات بهذه الطريقة تحقق شرط اللاتغاير (Invariance)؛ بمعنى أن التقديرات لا تختلف لأنها تستنبط مباشرة من مستوى القدرة للعينة.
- تقديرات (CML) لها خصائص مرغوبة في ظروف متعددة مثل الاتساق، وتقديرات درجة الصعوبة تتبع التوزيع الطبيعي، كما أن عملية التقدير فعّالة، وتُعد هذه التقديرات فريدة.
- وجود فائدة في استخدام لوغاريتم الأرجحية في فحص الفرضيات.
- الخطأ المعياري لمعالم الفقرات أكثر قابلية للتفسير كممثلة لأخطاء المعاينة.

عيوب أسلوب الأرجحية العظمى الشرطية:

- ينحصر استخدام هذه الطريقة في نموذج "راش" وتعميماته فقط، أما نماذج الاستجابة الأخرى الثنائية والثلاثية، والنماذج متعددة الأبعاد لا تقدّر معالمها بهذه الطريقة، لأن العلامة الكليّة لا تمثل إحصائيًا كافيًا للتقديرات.
- هناك فقدان لبعض المعلومات، فلا يوجد تقدير فعّال للفقرات والأفراد ذوي العلامات التامة (الكاملة والصفرية) .
- وجود مشكلات في عملية التقدير عندما يكون عدد الفقرات كبيرًا، فمشكلة الحسابات والأرقام تتطور في حالة الاختبارات الطويلة حيث يفشل التقدير غالباً عند(30) فقرة، لكن تغلبت البرامج الجديدة في (CML) فيما يتعلق باختبارات طويلة، إلا أن صعوبة التقدير ما زالت قائمة في الاختبارات الطويلة جدًا التي يزيد عدد فقراتها عن 100 فقرة.

بقي أن نقول أن هذه الطريقة (CML) شائعة الاستخدام في دول أوروبا الغربية، ويمكن استخدام برنامج الحاسوب (PML) في إجراء هذه الطريقة (علام، 2005).

(ج) طريقة الأرجحية العظمى الهامشية (Marginal Maximum Liklihood (MML: تعتمد هذه الطريقة على تعظيم عملية التقدير لمعالم الفقرات اعتمادًا على التوقع لكل من أعداد الطلبة الذين أجابوا على أي من الفقرات إجابة صحيحة، أو الذين تكون قدرتهم مساوية لأي من القدرات التي يمكن للطلبة الحصول عليها في توزيع القدرة المعتمده من جهة، وأعداد الطلبة في كل فئة من فئات القدرة معلومة التوزيع، وتسمى أيضاً "أسلوب تعظيم التوقع Expectation/Maximization (EM)" (Hambleton & Swaminthan,1985).

وتتطلب طريقة (MML) تحديد مستويات توزيع القدرة، وغالبًا يحدد كتوزيع سويّ، وإذا توفر حجم عينة كاف فإن التوزيع لا يتطلب أن يكون معروفًا مسبقًا، فقد يكون توزيع القدرة مقدرًا من توزيعات أخرى غير التوزيع السوي إذا أعتبر أنه الأنسب (Embretson & Reise, 2000) .

وأشار امبرسون ورايس (Embretson & Steven Reise, 2000) إلى عدد من المميزات والسلبيات لأسلوب (MML) هي:

المميزات:

- يمكن تطبيق هذه الطريقة على جميع نماذج (IRT) حتى مع النماذج متعددة الأبعاد.
- فعالة للاختبارات الطويلة والقصيرة.
- تقديرات (MML) للأخطاء المعيارية قد تفسر التقديرات الجيدة لتباين المعاينة المتوقع.
- التقديرات ذات قيمة (أو تتوافق) مع العلامات الكاملة، ولا يحدث فقدان للمعلومات من الأفراد، لأن الاستجابة الكاملة لها تكرار نسبي متوقع للفقرات؛ فالفقرات لا تحتاج إلى ترتيب وبخاصة إذا كان التوزيع لمعالم الفقرات قد تم تحديده بشكل جيد.

- يمكن تفسير البيانات بشكل جيد ومفيد لفحص الفرضيات ومؤشر المطابقة، فنسبة الأرجحية تستخدم لتحديد أي النماذج أكثر ملائمة، ويمكن تحديد نسبة الأرجحية لمربع كاي للاختبار.

السلبيات:

- صعوبة برمجة الخوارزمية، فبرامج الحاسوب يجب أن تكون فعالة حسابيًا ورقميًا.
- توزيع القدرة يجب أن يتم افتراضه مسبقًا؛ لذا فإن تقديرات المعالم مرهون بالملاءمة مع التوزيع المفترض، وهذه ليست سلبية كبيرة فقد يكون التوزيع المفترض غير سوي كأن يكون ملتوٍ أو مفلطح ويتم تكييفه باستخدام أوزان مناسبة للترجيح.

(د) نماذج تقدير أخرى:

هناك العديد من الطرق الأخرى للتقدير، حيث توجد طريقتان لهما تاريخ طويل هما:

- طريقة المربعات الصغرى التي تطور نماذج (IRT) والتي تحوي معلم التمييز، فقد طوّر (McDonald) المشار اليه في (التقي، 2009) طرق تقدير المعالم بالتحليل العاملي غير الخطي والذي تم توسيعه مباشرة للتوزيع الطبيعي لنماذج (IRT).
- طوَر (Chopin) التقدير المشروط المزدوج من نموذج "راش" كبديلٍ كافٍ لتقدير (JML) وهذه الطريقة تستخدم استجابات الأفراد على أزواج الفقرات مثل طريقة (CML) ومستوى القدرة يلغي تقدير صعوبة الفقرات (Embretson & Steven Reise, 2000).

تقييم حُسن مطابقة النموذج للبيانات:

تأتي أهمية مطابقة النموذج للبيانات Model – Data Fit للاستفادة من مميّزات النموذج عند استخدام برامج الحاسوب المختلفة في تقدير معالم النموذج استنادًا إلى البيانات الاختبارية،

حيث لا تتم الاستفادة من مميزات نماذج الاستجابة للفقرة إلا إذا كانت هذه المطابقة معقولة، وبدون ذلك لا تتحقق خاصية اللاتغاير Invariance لكل من معالم الفقرات وقدرة الأفراد، وهذا يتطلب التحقق من درجة اتساق البيانات الاختبارية مع الافتراضات التي يقوم عليها النموذج المستخدم، وأصبح معلومًا أن أهم الافتراضات التي تستند إليها جميع نماذج الاستجابة للفقرة هي: افتراض أحادية البعد، وافتراض الاستقلال الموضعي.

(أ) **تقييم افتراض أحادية البُعد:** قام هامبلتون وروجرز (Hambleton & Rogers,1988)، وهاتي (Hattie, 1985) بإجراء مراجعة شاملة للأساليب المستخدمة في تقييم أبعاد مصفوفة البيانات ومنها:

- التحقق من اتساق أنماط استجابات الأفراد للفقرات الاختبارية.
- نتائج برامج الحاسوب المتعلقة بتحليل المكونات الأساسية لمصفوفة البيانات الاختبارية، وتحديد عدد قيم الجذور الكامنة eigenvalues الأكبر من الواحد الصحيح، أو الرسم البياني لهذه القيم بعد ترتيبها تنازلياً من الأكبر إلى الأصغر، ودراسة هذه القيم لتحديد ما إذا كان هناك عامل أول أكثر تشبعاً أي أنه يفسر أكبر قدر من التباين المشترك في المصفوفة.
- نتائج برنامج التحليل العاملي غير الخطي الذي يتطلب إجراء مطابقة لنموذج التحليل العاملي غير الخطي لمصفوفة معاملات الارتباط الداخلية بين الفقرات، وفحص بواقي تحليل التغاير باستخدام برامج حاسوب خاصة مثل برنامج (NOHARM).
- التحقق من الفقرات التي من المحتمل مخالفتها لافتراضات نموذج الاستجابة المستخدم، وذلك بتقدير معالم هذه الفقرات ومقارنة نتائج التقدير هذه مع نتائج تقدير معالم المجموعة الكلية للفقرات، فإذا تبين من التمثيل البياني للقيم التقديرية لهذه المعالم في الحالتين أنه

خطيًا، والتشتت يقترب من الأخطاء المعيارية لهذه القيم، فإنه يمكن افتراض تحقق خاصية أحادية البعد لهذه البيانات الاختبارية. وقد أكد كل من (Hambleton & Rovinlli) المشار اليهما في (Hambelton & Rogers,1988) أن الأسلوب الثالث والمتعلق باستخدام التحليل العاملي غير الخطي قد نال الاهتمام من خبراء القياس كوسيلة ممكنة لتقييم أبعاد مصفوفة البيانات الاختبارية.

(ب) تقييم افتراض الاستقلال الموضعي: إن عدم تحقق هذه الخاصية يؤثر في تقدير معالم نماذج الاستجابة للفقرة، وربما يؤدي إلى قيم تقديرية أكبر من حقيقتها، ويؤدي أيضاً إلى أخطاء في تحديد عدد الأبعاد الكامنة في مصفوفة البيانات، هذا ما بينته الدراسات التجريبية لكل من ين (Yen,1993)، وستينبرغ، وثيسن (Steinberg & Thissen,1996). كذلك فإن العلاقة الآتية التي تحدد الاحتمال الشرطي للحصول على النمط من الاستجابات والتي تدل على اقتران الأرجحية لقيم القدرة (θ) لا تكون صحيحة في حال لم تتحقق الاستقلالية للإجابة عن الفقرات التي عددها L:

- $$L(\theta, x) = \prod_{i=1}^{L} P(\theta)^{u_i}(1 - P(\theta))^{1-u_i} \ldots\ldots\ldots\ldots [11]$$

حيث يدل الرمز u على الإجابة الثنائية، ويأخذ إحدى القيمتين (صفر) للإجابة الخاطئة و (واحد صحيح) للإجابة الصحيحة. واقترح ين (Yen,1993) الإحصائي (Q3) الذي تمثل قيمته عدم الاستقلالية المحلية (Local Dependence, LD) في أنماط الاستجابة التي يعطيها الأفراد. وتتلخص طريقة إيجاده بحيث يتم بعد تقدير كل من معالم الفقرات وقدرات الأفراد إيجاد التوقع لاستجابات فرد محدد على الفقرات ثم إيجاد الفرق بين توقع

استجابة الفرد والاستجابة الفعلية ليكون هذا الفرق المسمى بالباقي، ثم إيجاد معامل ارتباط قيم البواقي لأي زوج من الفقرات لجميع الأفراد. ويُذكر أن التوقع لمعامل الارتباط بين البواقي يساوي (N-1/1-) عند تحقق الاستقلالية، عندما يكون حجم العينة (N) كبيرًا. وتدل قيمة (Q3) الكبيرة على عدم الاستقلالية، ويُذكر أن قيمة هذا المقياس تقترب من الصفر في العينات كبيرة الحجم. وقد بيّن كثير من خبراء القياس المعاصر أنه لا يوجد أسلوب وحيد يمكن استخدامه في تقييم حُسن مطابقة النموذج للبيانات بحيث يكون كافيًا، وأكدوا على ضرورة الحصول على أدلة متعددة لتحقيق ملاءمة النموذج والحكم على ملاءمة النموذج الذي تم اختياره (Hambleton & Rogers,1988؛ Embretson & Steven Reise, 2000). ويمكن القول أن أفضل طريقة للتغلب على مشكلة عدم تحقق الاستقلال الموضعي هو تجنب حدوثها في مرحلة تصميم الاختبار، وتهيئة ظروف الموقف الاختباري (علام، 2005). ويُذكر أن (Hambleton & Swaminathan ,1985) ، قد عدّا افتراض الاستقلال الموضعي مكافئًا لافتراض أحادية البعد، وقد أكد على هذا كل من (Hambleton, Swaminathan & Rogers,1991)، حيث عدّا هذين الافتراضين متلازمين، أي أنه إذا تحقق الافتراض الأول فإن الافتراض الثاني يتحقق أيضاً.

(ج) تقييم التحرر من السرعة في الأداء (Speedeness): ويعني ذلك تطبيق الاختبار ليس كاختبار سرعة، بمعنى أن إخفاق الفرد في الإجابة على فقرات الاختبار يعود إلى انخفاض قدرته وليس بتأثير عامل السرعة على الإجابة، وتوفير الوقت الكافي بحيث ينهي جميع الطلبة الإجابة عن فقرات الاختبار ضمن الوقت المحدد للإجابة. وللتحقق من هذا الافتراض حسب (Hambleton & Swaminthan,1985) يتم باتباع الإجراءات الآتية:

- مقارنة التغاير في عدد الفقرات المحذوفة إلى التغاير في عدد الإجابات غير الصحيحة للفقرة.
- التأكيد على العلاقة بين درجات الاختبار في فترة زمنية محددة مع تطبيق الاختبار بدون تحديد فترة زمنية.
- التأكيد على النسبة المؤية للمفحوصين الذين أنهوا الاختبار والنسبة المؤية للمفحوصين الذين أنهوا 75% من فقرات الاختبار.
- التأكيد على عدد الفقرات التي أجاب عليها 80% من المفحوصين.

مطابقة الفقرات للنموذج

من المعلوم ان الاختبار الواحد يتكون من مجموعة من الفقرات من أنواع مختلفة، لذلك ينبغي التحقق من مطابقة هذه الفقرات لنموذج معيّن من نماذج الاستجابة للفقرة المختلفة وذلك بالتحقق من مطابقة كل فقرة للنموذج، خصوصًا إذا كان الاختبار يحتوي على أنواع مختلفة من الفقرات التي تتطلب نماذج مختلفة لتقدير معالمها. ويقصد بمطابقة الفقرات لنموذج معين هو إمكانية تفسير النموذج لكيفية استجابات الأفراد لإحدى الفقرات، أو التنبؤ بهذه الاستجابات (علام، 2005). وتتم عملية المطابقة من خلال أسلوبين، الأول يقوم على تمثيل ما يتم توقعه من خلال النموذج المعتمد مع ما يتم ملاحظته فعلياً من البيانات المشاهدة. أما الثاني فيتم من خلال إيجاد إحصائيات للتعبير عن الفرق بين ما هو مشاهد وما هو متوقع. وفي كلتا الحالتين يتم تجزئة قيم القدرة (θ) المقدرة بعد ترتيبها تصاعدياً إلى عدد من الفئات تتضمن كل فئة العدد ذاته من الأفراد؛ أي وضعهم في مجموعات بحسب درجتهم الكلية في الاختبار ثم إيجاد نسبة الذين أجابوا إجابة صحيحة وتعيين النقاط لجميع الفئات من خلال تحديد الإحداثي السيني على أساس قيم

الوسيط للقدرة في كل مجموعة، والإحداثي الصادي لنسبة اللذين أجابوا إجابة صحيحة ورسم المنحنى ومقارنته مع منحنى آخر هو منحنى خصائص الفقرة، ويتم مقارنة النقاط التي تمثل ما هو مشاهد مع ما هو متوقع، وكلما كان الفرق كبيرًا دلّ ذلك على عدم المطابقة أو عدم الموائمة (التقي، 2009). وفي الطريقة الأولى يمكن الحصول على قيمة الإحصائي مربع كاي(χ^2) كمؤشر لحسن مطابقة الاختبار ككل باستخدام الصيغة الآتية:

- $$\chi^2 = \sum_{j=1}^{k} f_j \frac{(P_{ij} - \acute{P}_{ij})^2}{P_{ij} Q_{ij}} \ldots\ldots\ldots\ldots\ldots [12]$$

حيث يشير الرمز (P_{ij}) إلى نسبة عدد الإجابات الصحيحة الفعلية (أو المشاهدة) للفقرة (i) بواسطة الفرد (j) ويشير ($\acute{P}_{ij}$) إلى النسبة المتوقعة لعدد الإجابات الصحيحة استناداً إلى القيم التقديرية للمعالم، ويشير الرمز (f_j) إلى عدد الأفراد الذين تقع قدرتهم عند مستوى معين، أما الرمز (Q_{ij}) فيشير إلى احتمال الإجابة الخطأ على الفقرة حيث ($Q_{ij} = 1 - P_{ij}$). أما عدد درجات الحرية للإحصائي (χ^2) فيساوي عدد مجموعات الأفراد (أو فئات القدرة) مطروحًا منه عدد معالم الفقرات التي تم تقديرها حيث يمكن استخدام هذه الصيغة في النماذج الأحادية أو الثنائية أو ثلاثية المعلمة. ومن الجدير ذكره أن (Hambleton & Swaminthan, 1985) قد أوضحا أن كثيرًا من قيم مربع كاي الدالة إحصائيًا تتأثر كثيرًا بكبر حجم عينة الأفراد، مما يضفي بعض الشك على هذه القيم وتفسيراتها. لذلك اقترحا استخدام القيم النسبية لمربع كاي عند تقييم حُسن مطابقة الفقرات بدلاً من الاعتماد على الدلالة الإحصائية. واقترح (Mislevy & Bock,2003) استخدام نسبة أرجحية مربع كاي (Likelihood Ratio) Chi – Square في حالة الاختبارات

التي تشتمل على أكثر من 20 فقرة كطريقة أخرى في تقييم حُسن مطابقة الفقرات للنموذج من خلال المعادلة الآتية التي يتم استخدامها في كل من البرمجية (BILOG3) (1990) والبرمجية (BILOG – MG) (1996) :

- $$\chi^2 = 2\sum_{g=1}^{G}\left[\mathbf{R_g}\log\frac{\mathbf{R_{ig}}}{\mathbf{N_g P(\theta_m)}} + (\mathbf{N_{ij}} - \mathbf{R_{ig}})\log\frac{\mathbf{R_{ig}}}{\mathbf{N_{ig}(1-P(\theta_m)}}\right]\ldots[13]$$

حيث يشير الرمز R_j إلى نسبة عدد الإجابات الصحيحة الفعلية (المشاهدة) عن الفقرة (i) داخل المجموعة g ضمن فئات القدرة بعد ترتيب قدرات الأفراد تصاعديًا والرمز (N_g) يشير إلى عدد الأفراد في المجموعة (g) من عدد فئات القدرة (G)، ونسبة عدد الإجابات الصحيحة المتوقعة ($P(\theta_m)$ والمستخرجة من منحنى خصائص الفقرة للقدرة التي تمثل الوسط للفئة g داخل تقدير مستوى قدرة المجموعة ويرمز لهذا المتوسط بالرمز (m).

ويشير (Embretson & Steven Reise,2000) إلى الأسباب الآتية والتي تحول دون مطابقة الفقرة للبيانات المشاهدة وهي:

- وجود أكثر من بُعد خلف إجابات الطلبة.
- عدم كفاية معالم الفقرة المعنية في النموذج المعتمد.
- منحنى خصائص الفقرة المعنية ليس متزايدًا.
- قد يتضمن المجتمع الذي تم فيه تقدير معالم الفقرات مجتمعًا فرعيًا غير متجانس مع المجتمع الكلي، وبالتالي يكون هناك عدم مطابقة بعض الأفراد للنموذج المعتمد.
- وجود خلل في صياغة الفقرة.

لكن قد يتساءل المرء عن الفائدة التي نحصل عليها عند مطابقة نماذج استجابة الفقرة لمجموعة بيانات الاختبار، ويجيب (Hambleton & Swaminthan,1985) عن هذا التساؤل، ويذكران الفوائد الآتية:

- تقدير قدرة المفحوص يمكن الحصول عليها على نفس مقياس القدرة بحيث يمكن مقارنتها بالرغم من أن المفحوصين قد أجابوا على مجموعات مختلفة من الفقرات الاختبارية.
- إحصائيات الفقرة التي يمكن الحصول عليها لا تعتمد على عينة المفحوصين التي تم استخدامها لمعايرة فقرات الاختبار.
- إضافة إلى مؤشر تقديرات القدرة على كل نقطة على مقياس القدرة التي تم الحصول عليها.

دالة معلومات الفقرات:

يُعدّ مفهوم دالة المعلومات من المفاهيم الأساسية في نظرية الاستجابة للفقرة، حيث تدل على العلاقة التي تربط الفقرة مع معلوماتها، وتشير إلى قدرة الفقرة على التمييز بين الأفراد عند مستوى قدرة معينة، وأنه لكل فقرة واختبار دوال تعرّف بشكل مستقل عن مجموعة المفحوصين، وتقدم كنظير للثبات في النظرية الكلاسيكية، إضافة إلى أنها تمثل الخطأ المعياري في القياس عند أي مستوى قدرة، ويعني ذلك أن دقة تقدير القدرة يمكن التعبير عنها بدالة المعلومات، أو منحنى المعلومات الذي يمثلها بيانياً. ولدالة المعلومات تطبيقات عديدة في بناء الاختبارات واختيار الفقرات وتقييم دقة القياس ومقارنة الاختبارات، وتحديد أوزان الدرجات ومقارنة طرق تدريج الدرجات (Hambleton & Swaminthan, 1985).

ولا تقتصر تطبيقات دالة المعلومات على نماذج الاستجابة للفقرة ثنائية الدرجة بل تعدتها إلى النماذج متعددة الأقسام، والنماذج متعددة الأبعاد، لذلك يُعد مفهوم دالة المعلومات من المفاهيم السيكومترية التي تتميز بالعمومية، فعندما نحصل على تقدير الأرجحية القصوى لمعلم القدرة؛ فإن تباين هذا التقدير يكون مقلوب دالة المعلومات المناظرة؛ وبالمثل عند تقدير معالم الفقرات بأسلوب الأرجحية العظمى فإن مصفوفة التباين – التغاير (Variance – Covarriance Matrix) للتقديرات تكون معكوساً لمصفوفة المعلومات (Information Matrix) المتعلقة بتقديرات معالم الفقرات، وتعتبر مصفوفة التباين المشترك لتقديرات معالم الفقرة مهمة عندما تتم مقارنة معالم الفقرة لمجموعتين من المفحوصين (Hambelton, Swaminathan & Rogers, 1991).

ونظرًا لأن دالة المعلومات تتباين بتباين مستوى القدرة، فقد اقترح خبراء القياس منحنيات لمعلومات الاختبارات يمكن أن تكون بديلاً لتقديرات الثبات المستخدمة في نماذج القياس الكلاسيكية وللأخطاء المعيارية للقياس المستخدمة في تفسير درجات الاختبارات؛ ويمكن إجراء تحويلات معينة على منحنى خصائص الفقرة (ICC) ليصبح منحنى معلومات الفقرة (Item Information Curve) (IIC)، وهذا المنحنى يبين كمية المعلومات السيكومترية التي تشتمل عليها فقرة اختبارية عند جميع نقاط متصل السمة الكامنة (علام، 2005).

يُذكر أنه توجد صيغ رياضية متعددة لدالة معلومات الفقرة تختلف باختلاف النموذج اللوجستي المعتمد. ولا يتسع المجال هنا لذكر تفاصيل هذه الصيغ، لكن من المفيد ذكر الملاحظات التي أبداها كل من (Hambelton & Swaminathan, 1985؛ علام،2005) وهي:

- يكون مقدار المعلومات المستمدة من الفقرة أكبر ما يمكن حول معلم صعوبة الفقرة (b_i)، لذلك ينبغي أن تتمركز صعوبة كل فقرة حول قدرة الفرد لكي تسهم الفقرة بأكبر قدر من المعلومات فيما يتعلق بمستوى قدرة الفرد. أما فيما يتعلق بالنموذج ثلاثي المعلمة فإن أكبر قدر من المعلومات يحدث عند مستوى أقل قليلاً من قيمة معلمة الصعوبة (b_i)، ذلك لاحتمال تأثير معلمة التخمين (c_i).
- هناك علاقة طردية بين ميل منحنى المعلومات ومقدار المعلومات التي تسهم بها الفقرة، ومن المعلوم أن الميل يمثل معلمة التمييز، فكلما زادت قيمة تمييز الفقرة زاد مقدار المعلومات التي تسهم بها الفقرة حول معلمة الصعوبة عند كل مستوى من مستويات القدرة، والعكس إذا انخفضت معلمة التمييز فإن منحنى المعلومات يصبح أكثر انبساطًا وانتشارًا ويقل معه ميل المنحنى مما يقلل من المعلومات التي تسهم بها الفقرة.
- تأخذ في العادة منحنيات دوال المعلومات للفقرات شكل الناقوس أو الجرس، ويكون أقصى حد من المعلومات التي نحصل عليها عند معلمة الصعوبة على متصل القدرة لكل من النموذجين أحادي وثنائي المعلمة؛ بينما نحصل على أقصى معلومات في حالة النموذج ثلاثي المعلمة وفق الصيغة الرياضية الآتية :

$$- \quad I_i(\theta_{\max}) = b_i + \frac{1}{D\alpha_i}\ln\left[\frac{1}{2}+\frac{1}{2}\sqrt{1+8c_i}\right] \ldots\ldots\ldots\ldots\ldots\ldots [14]$$

والمقدار D ثابت قيمته (1.7) يستخدم للاقتراب من المنحنى الاعتدالي، وعند عدم وجود التخمين (c=0) فإن : $I_i(\theta_{max}) = b_i$.

- تكون أقصى قيمة للمعلومات ثابتة في النموذج أحادي المعلمة؛ بينما تتناسب أقصى قيمة للمعلومات في النموذج ثنائي المعلمة طرديًا مع مربع القيمة التقديرية لمعلمة التمييز للفقرة (α_i).

- نظراً لتأثر القيمة القصوى لدالة المعلومات للفقرة بقيم معالم الفقرة، والموضع الذي نحصل فيه على هذه القيمة القصوى فإنه يمكن انتقاء الفقرات التي تسهم في دقة القياس على متصل القدرة مما يساعد في تصميم الاختبارات، كما يحدث في الاختبارات المحوسبة.
- يمكن إيجاد مقدار المعلومات التي تسهم بها كل فقرة عند كل نقطة على متصل القدرة.
- كلما قلّ التباين المتعلق بفقرة ما؛ يزداد مقدار المعلومات التي تسهم بها هذه الفقرة.
- يتأثر مقدار المعلومات بشكل عام بنوعية الفقرات الاختبارية وعددها.

يؤكد (Hambelton & cook ,1983) على أن عدد المفحوصين وعدد فقرات الاختبار يؤثران على استقرار ودقة تقدير معالم الفقرات، والذي بدوره يؤثر في مقدار المعلومات التي تقدمها الفقرة (دالة المعلومات)؛ كما أنه من الممكن أن يكون لعدد فقرات الاختبار تأثير على صدق وثبات الاختبار. وبيّن Lord (1980)، أن طول الاختبار وحجم عينة الأفراد المفحوصين يؤثران على جودة تقدير المعالم، ومن هنا جاءت فكرة هذه الأطروحة في دراسة أثر حجم العينة وعدد فقرات الاختبار ونوعها على دقة تقدير معالم الفقرة والقدرة ودالة المعلومات.

- اسهام كل فقرة يكون مستقلاً عن اسهامات الفقرات الأخرى التي يتكون منها الاختبار؛ لذا لا يعتمد مقدار المعلومات $I(\theta)$ على تركيبة معينة من فقرات الاختبار.
- يرتبط مقدار المعلومات التي تسهم بها مجموعة من الفقرات عند مستوى قدرة معيّن ارتباطًا عكسيًا بالخطأ المتعلق بتقدير القدرة؛ وتُعطى العلاقة بين الخطأ المعياري لتقدير القدرة $SE(\theta)$ ودالة المعلومات حسب الصيغة الآتية:

$$\boldsymbol{SE(\theta) = \frac{1}{\sqrt{I(\theta)}}} \dots\dots\dots\dots\dots\dots [15] \quad -$$

لذلك يتم استبعاد الفقرات التي تكون الأخطاء المعيارية للتقدير بالنسبة لها كبيرًا عند بناء الاختبار.

دالة معلومات الاختبار:

تعتمد دالة معلومات الاختبار على تجميع إسهامات كل فقرة من فقرات الاختبار من المعلومات عند مستوى قدرة معيّن حسب العلاقة:

$$\boldsymbol{I(\theta) = \sum_{i=1}^{n} I_i\,(\theta)} \ldots\ldots\ldots\ldots\ldots\ldots..[16] \quad -$$

$$\boldsymbol{I(\theta) = \sum_{i=1}^{n} \frac{[\acute{P}(\theta)]^2}{[P_i(\theta)][1-P_i(\theta)]}} \ldots\ldots\ldots\ldots\ldots\ldots..[17] \quad -$$

حيث تمثل $I(\theta)$ دالة فقرات الاختبار؛ وتمثل $I_i(\theta)$ دالة معلومات الفقرة؛ ويشير الرمز $\acute{P}_i(\theta)$ إلى المشتقة الأولى لدالة خصائص الفقرة عند مستوى قدرة معين للفقرات التي عددها (n).

دالة الفعالية النسبية:

ويطلق عليها أيضاً الكفاءة النسبية (Relative Efficiency). إن بعض تطبيقات القياس النفسي والتربوي تتطلب مقارنة الفعالية النسبية لاختبارين في قياسهما قدرة معينة عند نقاط مختلفة على متصل القدرة؛ ويمكن إجراء هذه المقارنات بين دوال المعلومات بحساب الفعالية النسبية لأحد الاختبارات مقارنة باختبار آخر، ويُعبّر عن ذلك رياضيًا بالعلاقة:

$$\boldsymbol{RE(\theta) = \frac{I_A(\theta)}{I_B(\theta)}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots [18] \quad -$$

حيث $RE(\theta)$: الفعالية النسبية (أو الكفاءة النسبية).

$I_A(\theta)$: دالة المعلومات للإختبار A

$I_B(\theta)$: دالة المعلومات للإختبار B

وتكون الدالتان معرفتين على متصل قدرة مشترك (θ).

من خلال هذه العلاقة يمكن مقارنة الاختبارين وإضافة فقرات مثلاً ليكونا متماثلين ويؤدي ذلك إلى دقة في القياس، ويتم حساب الكفاءة النسبية عند كل نقطة من متصل القدرة كما هو

الحال في دالة المعلومات؛ غير أن دالة المعلومات لا يمكن تفسيرها في عبارات مطلقة ما لم يتم تعريف تدريج قياس القدرة ($\boldsymbol{\theta}-\mathbf{metric}$)، وأي تحويل لنظام التدريج يؤثر في هذه الدالة، بينما لا يؤثر ذلك في دالة الفعالية النسبية، وهذا يضفي أهمية على هذه الدالة الأخيرة ويثري استخداماتها في المقارنة بين الاختبارات التقليدية والاختبارات المحوسبة والتكيفية، وغيرها من التطبيقات السيكومترية (Hambleton & Swaminathan, 1985). فعلى سبيل المثال، إذا كانت $I_A(\theta)$ = 60 ، $I_B(\theta)$ = 40 مثلاً، فإن $RE(\theta)=\frac{60}{40}=1.5$. وهذا يعني أن الاختبار (A) يُعد كما لو كان أطول من الاختبار *(B)* بمقدار 60%. لذلك فإن الاختبار *(B)* يحتاج إلى زيادة في طوله بمقدار 50% بإضافة فقرات مماثلة للفقرات التي يشتمل عليها بالفعل لكي يؤدي إلى نفس دقة الاختبار (A) عند مستوى القدرة (θ) ؛ وبالمثل يمكن تقليل عدد فقرات الاختبار (A) بنسبة 40% ويظل محتفظًا بدقته في تقدير القدرة (θ) كالاختبار(B). وفي كلا الحالتين يفترض أن الفقرات التي تضاف أو تحذف يكون لها خصائص إحصائية مماثلة لخصائص بقية الفقرات. وبشكل عام يؤكد (Crocker & Algina, 1986) أنه لأي درجة سمة كامنة عندما تكون (RE(θ,A,B)>1) فإن A يكون قياس أفضل، وبصورة مشابهة عندما تكون درجة السمة الكامنة (RE(θ,A,B)<1) فإن *B* يكون قياس أفضل للسمة الكامنة.

أدوات القياس والتقويم:

تتعدد أنواع التقويم وأغراضه ومجالاته ؛ فقد يعتمد المعلم الاختبارات بمختلف أنواعها في جمع المعلومات الضرورية إلى جانب أدوات أخرى مثل التقارير، والتعيينات البيتية، والتقديرات الذاتية وغيرها. وقد صنف هبكنز وأنتس (Hopkins and Antes,1978) هذه الأدوات إلى فئتين هما:

- فئة الأدوات التي تندرج تحت اسم الاختبارات، وتشمل الاختبارات المقننة أو المنشورة؛ والاختبارات التي يُعدّها المعلم؛ وهي في الأغلب اختبارات تحصيلية.
- فئة الأدوات التي لا تندرج تحت اسم الاختبارات وتشمل الأدوات التي تصمم لقياس بعض الأهداف التي لا تقاس بالاختبارات بالمعنى التقليدي، وتتعلق بأداء معقد نسبيًا وتشمل التقارير، والأدوات البحثية والمقالات والنماذج، والمشاريع، والسجلات الصحفية.

تعريف الاختبار:

تتكرر كلمة "امتحان" أو "اختبار" أكثر من غيرها في الأوساط التربوية لأنها تعد من أكثر أدوات القياس والتقويم استخدامًا من قبل المعلمين في كافة المراحل. وتُعد كلمة مقياس أعم من كلمة اختبار أو امتحان. ويُعرّف الاختبار بشكل عام، بأنه أداة قياس يتم إعدادها وفق طريقة منظمة من عدة خطوات تتضمن مجموعة من الإجراءات التي تخضع لشروط وقواعد محددة بغرض تحديد درجة امتلاك الفرد لسمة معينة من خلال إجاباته عن عينة من المثيرات التي تمثل السمة المرغوب قياسها (عودة، 2004). أي أن ما نقيسه هو قدرات الفرد أو السمات التي تمثل مجموعة من السلوكات المترابطة، حيث تختلف درجة امتلاك الفرد لسمة معينة عن أي سمة أخرى، كما تختلف درجة امتلاك السمة نفسها من فرد إلى آخر، ويؤمل أن تكشف نتائج القياس عن هذه الفروق.

ويتعامل المعلم مع أكثر من نوع من السمات، فهو يتعامل مع التحصيل بصفة رئيسية؛ وهذا ما يشار اليه عادة بالقياس التربوي؛ وهناك سمات أخرى مثل القلق والذكاء أو سمات شخصية تتعلق بالقيم والميول والاتجاهات؛ أو ما يشار اليه عادة بالقياس النفسي؛ ويتعامل مع نوع ثالث من السمات مثل الطول والوزن والقدرة السمعية والقدرة البصرية أو ما يشار إليه عادة بالقياس الفيزيائي؛

وعليه يكون القياس الصفي هو قياس نفسي – تربوي بالدرجة الأولى (Hopkins and Antes, 1978). ويقصد بالمثيرات هنا فقرات أو أسئلة الاختبار، وإجابات الفرد عن هذه الفقرات يشكل السلوك الدال عليها، وبما أن الاختبار أداة قياس، وأن القياس كمي فإن تعريف الاختبار يتضمن التقدير الكمي لدرجة امتلاك الفرد للسمة. وبناء على ما سبق يمكن تعريف الاختبار التحصيلي بأنه طريقة منظمة لتحديد مستوى تحصيل الطالب لمعلومات ومهارات في مادة دراسية كان قد تم تعلمها مسبقا بصفة رسمية، من خلال إجاباته عن عينة من الأسئلة أو الفقرات التي تمثل محتوى المادة الدراسية، وتشمل اختبارات التحصيل العام أي في عدة مباحث مجتمعة، أو اختبارات التحصيل الخاص في مبحث معين (عودة، 2004).

أنواع الفقرات:

يصنّف المشتغلون بالقياس الاختبارات حسب شكل الفقرة Item – Format حيث توجد أنواع مختلفة من الفقرات التي يمكن أن يستخدمها المعلم في اختباراته تنقسم إلى فئتين رئيسيتين هما:

- فئة الفقرات ذات الإجابة المنتقاة (الموضوعية) وتشمل:
  - الصواب والخطأ.
  - الاختيار من متعدد.
  - المطابقة أو المزاوجة.
- فئة الفقرات ذات الإجابة المصوغة (المقالية) وتشمل:
  - التكميل.
  - الإجابة القصيرة.

- الإنشائية المحددة.
- الإنشائية المفتوحة.

وإذا أردنا الحديث عن هذين النوعين الرئيسين من الفقرات؛ فإن فئة الفقرات من النوع الأول والتي يُطلق عليها الفقرات ذات الإجابة المنتقاة (الموضوعية) تتطلب من المفحوص أن يختار الإجابة الصحيحة أو الإجابة الأصح من بين مجموعة من البدائل. والفئة الثانية وتسمى فئة الفقرات ذات الإجابة المصوغة أو الاسئلة المقالية (الاسئلة الانشائية) تتطلب من المفحوص صياغة الإجابة بلغته الخاصة. ويتوقف استخدام أي من هذه الأنواع من الفقرات على الهدف من عملية التقويم؛ حيث تُستخدم الفقرات الموضوعية أو ذات الإجابة المنتقاة غالباً لقياس مدى امتلاك الطالب لمهارات التذكر والفهم والتطبيق ولتغطية عينة أوسع من المادة المراد معرفة درجة اتقان الطالب لها. أما الأسئلة المقالية أو الانشائية فغالباً ما تُستخدم عندما يكون الهدف هو قياس القدرات التنظيمية أو قدرة الطالب على التعبير وإبداء الرأي؛ كما يتحكم في اختيار الشكل المناسب للفقرات عوامل عدة منها: طبيعة المادة الدراسية وعدد الطلبة وظروف التطبيق والإمكانات المدرسية، والأهم من ذلك هي مهارة المعلم وخبرته في وضع الأسئلة والاختبارات بشكل عام (عودة، 2004).

الدراسات السابقة:

أثار موضوع دقة تقدير معالم الفقرة والقدرة اهتمام العديد من الباحثين، وكان موضوع الكشف عن أثر حجم العينة وطول الاختبار ونوع فقراته في دقة تقدير معالم الفقرة وقدرة الفرد مدار بحث للعديد من الدراسات؛ يتناول هذا الجزء مراجعة للدراسات ذات الصلة بموضوع هذه الدراسة

التي تم الاطلاع عليها من قبل الباحث، وقد تم عرض هذه الدراسات وفقاً للترتيب الزمني بدءاً بالأقدم فالأحدث، ومن ثم وضع الباحث تعقيباً عليها يبرز ما تتميز به الدراسة الحالية عن سابقاتها

هدف بارنس ووايز (Barnes & Wise,1991) في دراستهما معرفة تأثير عدد فقرات الاختبار وحجم العينة على تقدير قدرة المفحوصين وتقدير معلمة الصعوبة ؛ وذلك باستخدام نموذج راش أحادي المعلمة من نظرية الاستجابة للفقرة، واستخدما في ذلك عينات صغيرة الحجم(50،100،200) مفحوص، وتطبيق اختبارين: الأول مكون من (25 فقرة) والآخر مكون من (50 فقرة)، حيث أشارت نتائج هذه الدراسة إلى عدم وجود تأثير لحجم العينة على تقدير قدرة المفحوصين، بل يتأثر التقدير بعدد الفقرات، وبالنسبة لمعلمة الصعوبة كان التغير في تقديرها بسيطا عند تغير حجم العينة .

في دراسة قام بها ليم وهون وهاردويل وميشيل (Lim, Hoon, Hardwell & Michael,1997) تناولت أثر عدد الفقرات وعدد المفحوصين على دقة تقدير المعالم في النموذج اللوجستي الثنائي من خلال إجراء تحليل لسبع دراسات سابقة، وذلك لتقديم معلومات عن نظرية الاستجابة للفقرة، بيّنت النتائج أن أكثر من 63% من التغيّر في المؤشرات المستخدمة لقياس دقة تقديرات المعالم تُعزى إلى هذين العاملين.

وفي دراسة أجراها الدرابيع (2001) هدفت إلى التحقق من فعالية النموذج اللوغاريتمي ذي المعلمة الواحدة "نموذج راش" في تقدير قدرة الفرد ومعلمة صعوبة الفقرة باختلاف حجم العينة وطول الاختبار، فقد استخدم حجم عينة (50،100،500) وطول اختبار (25،50،300) فقرة، وتم توليد بيانات ثنائية الإستجابة. وأشارت نتائج الدراسة إلى وجود فروق ذات دلالة إحصائية لتفاعل كل من حجم العينة وطول الاختبار على دقة تقدير قدرة الفرد، ووجود فروق ذات دلالة

إحصائية في دقة تقدير قدرة الفرد ترجع إلى متغير طول الاختبار وحده، في حين لم تُظهر النتائج وجود فروق ذات دلالة إحصائية تُعزى لمتغيّر حجم العينة. وفيما يتعلق بدقة تقدير معامل صعوبة الفقرة، فقد أشارت النتائج إلى وجود فروق ذات دلالة إحصائية لتفاعل كل من حجم العينة وطول الاختبار، وكان ثمة فروق ذات دلالة إحصائية في دقة تقديرمعامل صعوبة الفقرة ترجع إلى كل من حجم العينة وطول الاختبار كل على حده.

أجرى بيلتون (Pelton,2002) دراسة استخدم فيها النظرية الكلاسيكية ونماذج نظرية الاستجابة للفقرة واسلوب المحاكاة ،وذلك بهدف مقارنة الدقة والاستقرار في تقدير معلمة صعوبة الفقرة ومعلمة قدرة الأفراد، حيث توصل إلى أن تقديرات القدرة تتباين حسب كمية المعلومات المتوفرة من البيانات المولدة، وتتأثر هذه التقديرات بكل من أحادية البعد ودرجة التخمين ومدى التباين في صعوبة الفقرات، وتوصل في دراسته هذه أيضاً إلى أن النموذج اللوجيستي ثنائي المعلمة يقدم تقديرات دقيقة لمعلمة الصعوبة أكثر منها في كل من نموذج راش والنموذج ثلاثي المعلمة والنظرية الكلاسيكية، وذلك عند وجود حجم عينة بلغ (999) مفحوصا. وعدد من فقرات الاختبار بلغ (33 فقرة) مع وجود تخمين قليل في الإجابة .

وأجرى عبابنة (2004) دراسة بعنوان "أثر حجم العينة وطرق انتقائها وعدد الفقرات وطريقة انتقائها على دقة تقدير معالم الفقرة والقدرة لاختبار قدرة عقلية باستخدام نظرية الاستجابة للفقرة". وطبق الباحث الاختبار المؤلف من أربعة اختبارات فرعية هي: اختبار المفردات، واختبار المتشابهات، واختبار المتضادات، واختبار الحساب على عينة مكوّنة من (1000) طالب وطالبة من الصف السابع الأساسي يدرسون في مدارس وزارة التربية والتعليم التابعة لمديرية التربية والتعليم لمنطقة إربد الأولى، مستخدمًا برمجية (Bilog 3.11) لتقدير القدرة ومعالم الفقرات والأخطاء

المعيارية في التقدير، والتحقق من مطابقة البيانات للنموذج اللوجيستي ثلاثي المعلمة. وقد توصل الباحث إلى النتائج الآتية:

(أ) تزداد الدقة في تقدير معالم الفقرة بزيادة حجم عينة المفحوصين.

(ب) تتأثر الدقة في تقدير القدرة للمفحوصين بوضع مطابقة البيانات للنموذج ويكون عندئذ من غير المؤكد الحصول على دقة أكبر عند زيادة حجوم العينات المستخدمة في المعايير عند مستوى معين.

(ج) تزداد الدقة في تقدير معلم التمييز بزيادة تباين قدرة المفحوصين، كما تزداد الدقة في تقدير معلم صعوبة الفقرة عندما يكون توافق بين مدى قدرة المفحوصين ومدى صعوبة الفقرات. وكذلك تزداد الدقة في تقدير معلمة التمييز عند استخدام عينة من ذوي القدرة المتدنية في معايرة الفقرات.

(د) تزداد الدقة في تقدير معلمة القدرة بزيادة عدد فقرات الاختبار، أو نسبتها إلى الاختبار الكلي.

(هـ) هناك عدم استقرار في تقديرات معالم الفقرة عند استخدام عينات متغايرة القدرة.

(و) تقديرات معلمة القدرة أكثر دقة عند استخدام الطريقة المعيارية standard method في اختيار الفقرات مقارنة بالطريقة العشوائية random method.

وقام جلاس (Glass, 2005) بدراسة هدفت إلى معرفة أثر حجم العينة وعدد الفقرات في دقة تقدير معلمة قدرة الفرد وفق طريقة بييز التي تعتمد أسلوب تعظيم الاقتران، وذلك من خلال توليد بيانات ثنائية التدريج بأحجام عينات مقدارها (500، 1000، 2000) وفق النموذج الثلاثي المعلمة، وفقرات بعدد (200، 440) فقرة تم توليدها من بنك للأسئلة، وقد تم اعتماد المتوسط الحسابي للأخطاء المعيارية لمعلمة قدرة الفرد عند مستويات مختارة من القدرة موزعة على أطراف

متصل القدرة (2، 1، 0، −1، −2) لوجيت، عند أحجام العينات المختلفة والفقرات. وقد أظهرت نتائج الدراسة أنه بزيادة عدد الفقرات عند أحجام العينات (500 ،2000) تقل الأخطاء المعيارية وهذا بدوره يؤدي إلى الزيادة في دقة التقدير. أما عند زيادة عدد الفقرات عند حجم العينة (1000) تقل دقة التقدير.

هدفت دراسة المصري (2009) إلى الكشف عن تأثير شكل صياغة فقرات الاختبار على كل من تقدير قدرات المفحوصين والخطأ المعياري وخصائص الفقرات السيكومترية (الصعوبة والتمييز) والخطأ المعياري في تقدير كل منهما، واستخدم لذلك شكلين من الفقرات هما الاختيار من متعدد والتكميل، وبنى اختبار مكون من (40 فقرة) في الثقافة الحاسوبية لطلبة الحادي عشر، حيث صاغ الفقرات على شكل اختيار من متعدد بأربعة بدائل ثم صاغ تلك الفقرات على شكل أسئلة تكميل، ثم طبق الاختبار بشكليه على عينتين عشوائيتين حيث تم توزيع شكلي الاختبار على الطلبة في كل مدرسة بشكل عشوائي لكل شعبة من طلاب الحادي عشر حيث أظهرت النتائج أنه لا يوجد فرق ذات دلالة (α =0.05) في تقدير قدرات الأفراد تعزى إلى شكل فقرات الاختبار، ووجود فرق ذات دلالة (α =0.05) في تقدير صعوبة الفقرات يعزى إلى شكل فقرات الاختبار ولصالح فقرات التكميل، أي أن فقرات التكميل كانت أكثر صعوبة من فقرات الاختيار من متعدد، ووجود فروق ذات دلالة إحصائية (α =0.05) في تقدير تمييز الفقرات تعزى إلى شكل فقرات الاختبار ولصالح فقرات التكميل، أي أن فقرات التكميل كانت أكثر تمييزا من فقرات الاختيار من متعدد.

وأجرى الثوابية (2010) دراسة هدفت إلى استقصاء أثر حجم العينة في تقدير معلمة صعوبة الفقرة والخطأ المعياري في تقديرها باستخدام نظرية الاستجابة للفقرة. ولتحقيق هذا الهدف طبق الباحث اختبارًا تحصيليًا في الرياضيات للصف العاشر الأساسي تكون من (80) فقرة من نوع الاختيار من متعدد على عينات عشوائية طبيعية تراوح حجمها ما بين (200) إلى (11292) طالبًا وطالبة، مستخدما برمجية (Bilog – MG)، وقد توصل الباحث إلى النتائج الآتية:

أ– تزداد قيمة معلمة صعوبة الفقرة بزيادة حجم العينة حيث قفز متوسط صعوبة فقرات الاختبار من (0.31) لوجيت عند حجم عينة (200) طالب وطالبة إلى (1.1) لوجيت عندما أصبح حجم العينة (11292) طالبًا وطالبة.

ب– يتناقص الخطأ المعياري (SE) في تقدير معلمة الصعوبة بزيادة عدد أفراد العينة حيث بلغ (0.32) عندما كان حجم العينة (200) وتناقص ليصبح (0.07) لوجيت عندما أصبح حجم العينة (11292).

وأجرى العنزي (2010) دراسة هدف من خلالها الى التعرف على أثر شكل فقرة الاختبار على معالم الفقرة وخصائص الاختبار السيكومترية وفق نظرية الاستجابة للفقرة. حيث بنى اختبارين في مادة الرياضيات للصف الثالث الأساسي في السعودية تكونت فقرات الاختبار الأول من نوع الاختيار من متعدد، والثاني كانت فقراته من نوع التكميل وبعدد اجمالي بلغ (35 فقرة) لكل اختبار. وتم تطبيق الاختبارين على (926 طالب)، وتوصل في نهاية دراسته الى عدم وجود أثر لشكل الفقرة في كل من معامليَ الصعوبة والتمييز ودالة المعلومات، ووجود أثر لشكل الفقرة في كل من الخطأ المعياري في تقدير صعوبة الفقرات لصالح اختبار التكميل، ووجود أثر لشكل الفقرة في الانحراف المعياري لخطأ تقدير تمييز الفقرات ودالة المعلومات لصالح اختبار الاختيار من متعدد.

وأشارت النتائج لعدم وجود فرق بين المتوسطات الحسابية لخطأ تقدير القدرة، وكذلك عدم وجود فروق بين معامليَ الثبات لشكليَ الاختبار.

وهدفت دراسة الشريفين (2012) إلى الكشف عن أثر طريقة تقدير معالم الفقرة وقدرات الأفراد على قيم معالم الفقرة، والخصائص السيكومترية للاختبار في ضوء تغيّر حجم العينة، حيث تم بناء اختبار تحصيلي في الفيزياء من نوع الاختيار من متعدد مكون من (33) فقرة، وطبق على عينة الدراسة المكونة من 1000 طالبًا وطالبة من طلبة الثاني الثانوي العلمي. وبينت النتائج وجود فروق ذات دلالة إحصائية بين متوسطات الأخطاء المعيارية لتقديرات معالم الفقرات يُعزى للتفاعل بين طريقة التقدير وحجم العينة، في حين لم تظهر فروق ذات دلالة إحصائية تُعزى لحجم العينة أو طريقة التقدير. كما أشارت النتائج إلى وجود فروق ذات دلالة إحصائية بين متوسطات الأخطاء المعيارية لتقديرات قدرة الأفراد تُعزى لمتغير حجم العينة وللتفاعل بين طريقة التقدير وحجم العينة. كذلك بينت النتائج عدم وجود فروق ذات دلالة إحصائية بين معاملات الثبات المقدّرة وفق نظرية الاستجابة للفقرة عند أحجام العينة المختلفة (100، 500، 1000). وأشارت النتائج إلى أن دقة تقديرات معلمة القدرة تزداد في حالة عينة الأفراد ذوي القدرة العالية وعينة الأفراد ذوي القدرة المتدنية عند استخدام طريقة بييز التوقع (EAP)، في حين تزداد الدقة عند مستويات الأفراد ذوي القدرة المتوسطة باستخدام طريقة الأرجحية العظمى (MLE) بغض النظر عن حجم العينة.

بينما توصل الحجيلي (2012) في دراسته التي هدفت إلى معرفة تأثير طول الاختبار وحجم العينة في دقة تقدير كل من الصعوبة والقدرة وفاعلية معادلة الاختبارات بوجود الأداء التفاضلي للفقرة مستخدما نموذج راش الهرمي، وتحت شروط مختلفة للأداء التفاضلي، مستخدما حجم عينة (800،1600مفحوصا) وتوليد استجابة لاختبارين من (12،20 فقرة) مع وجود فقرات

مشتركة هي (3 و 5 فقرات) حيث أظهرت النتائج وجود ارتباطات قوية ودالة بين تقديرات الصعوبة والقدرة، وتوصل أيضا إلى أن قيم الخطأ المعياري تختلف بشكل دال عند طول الاختبار وحجم العينة والتفاعل بينهما .

وهدفت دراسة الزبون (2013) إلى معرفة أثر حجم العينة على تقدير دالة المعلومات للاختبار والخطأ المعياري في تقديرها باستخدام نظرية الاستجابة للفقرة، حيث وزع عينة الدراسة المكونة من (7500) طالبًا وطالبة من طلبة الصف الثامن الأساسي إلى خمس مجموعات بواقع (500، 1000، 1500، 2000، 2500) طالبًا وطالبة على التوالي، واستخدم استجاباتهم على الاختبار الوطني لضبط نوعية التعليم لمادة الرياضيات المكون من (40) فقرة من نوع الاختيار من متعدد، وباستخدام برنامج (Biolg– mg3)، والبرنامج الإحصائي (SPSS). وأظهرت النتائج أن مقدار تقديرات دالة المعلومات يتغير بتغير حجم العينة بطريقة طردية، إذ يزداد بزيادة حجم العينة، وأن الخطأ المعياري في تقدير دالة المعلومات يتغير بتغير حجم العينة بطريقة عكسية؛ إذ يتناقص بزيادة حجم العينة.

وتوصّل كينج بينج وتشرس ويدون (QingPing & Chris Wheadon,2013) إلى أن حجم العينة المستخدمة لمعايرة الفقرات أو تقدير معالم الفقرات في نظرية الاستجابة للفقرة (IRT) يعد من العوامل الرئيسية التي تؤثر على دقة تقدير معالم الفقرة، حيث تم استخدام اختبار تحصيلي في الرياضيات مكون من فقرات ثنائية الاستجابة وأخرى متعددة الإجابة، وتم توليد بيانات كبيرة، واستخدم نموذج التقدير الجزئي (Partial Credit Model) للفقرات المتعددة الاستجابة، والنموذج أحادي المعلمة "نموذج راش" للفقرات ثنائية الاستجابة. وقد توصل الباحثان إلى أن دقة تقدير معالم الفقرات يتأثر أيضاً بعدد فئات الإجابة وتوزيع درجة الفئات ضمن الفقرات، كما توصلا إلى

أن أخطاء القياس المرتبطة بتقدير معالم الفقرات تكون أعلى في حالة نموذج التقدير الجزئي منه في نموذج "راش".

تعقيب على الدراسات السابقة:

من خلال الدراسات التي تم عرضها سابقاُ يمكن استخلاص بعض التوجهات في النتائج التي أمكن التوصل اليها فيما يتعلق بدقة تقدير معالم الفقرة والقدرة وذلك على النحوالآتي:

1) إن دقة تقدير المعالم في نماذج نظرية الاستجابة للفقرة تتأثر بالعديد من العوامل أو المتغيرات مثل نوع فقرات الاختبار وحجم عينة المفحوصين، وطريقة انتقائهم، والطريقة المتبعة في التقدير، وطريقة انتقاء الفقرات؛ والطريقة المستخدمة في تقدير المعالم والبرنامج الحاسوبي المستخدم.

2) بيَنت الدراسات السابقة أهمية نظرية الاستجابة للفقرة وتوجه الكثير من الباحثين اليها في قياس دقة تقدير المعالم.

3) استخدمت بعض الدراسات السابقة الخطأ المعياري في التقدير (SEE) كمحك لتقييم دقة تقدير المعالم في نماذج نظرية الاستجابة للفقرة، بينما استخدم البعض الآخر الجذر التربيعي لمتوسط مربعات الأخطاء (RMSE)، واستخدمت دراسات أخرى محكات أخرى مختلفة.

4) تتميَز هذه الدراسة عن غيرها من الدراسات بأنها استهدفت الكشف عن أثر المتغيرات الثلاث (حجم العينة، وطول الاختبار، ونوع أوشكل الفقرات) وتفاعلاتها الثنائية والثلاثية على دقة تقدير معالم الفقرة وقدرة الأفراد ودالة معلومات الاختبار، ذلك من خلال استخدام بيانات واقعية وحقيقية يتم جمعها من الميدان مباشرة وعدم الاعتماد على بيانات مولّدة، واستخدم في سبيل ذلك اختبار تحصيلي في مادة الرياضيات للصف الثامن الأساسي كأداة للدراسة.

الفصل الثالث

الطريقة والإجراءات

يتناول هذا الفصل وصفًا للطريقة والإجراءات التي تم استخدامها في هذه الدراسة من أجل تحقيق هدف الدراسة، وهو قياس أثر حجم العينة وعدد الفقرات ونوعها على دقة تقدير معالم الفقرة والقدرة ودالة المعلومات لاختبار تحصيلي في الرياضيات للصف الثامن الأساسي.

مجتمع الدراسة:

تكوَن مجتمع الدراسة من جميع طلبة الصف الثامن الأساسي الذكور والإناث في مدارس مديرية تربية لواء قصبة عمان (عمان الأولى سابقاً) للعام الدراسي 2016/2015 - الفصل الدراسي الأول - والبالغ عددهم (7404) طالبًا وطالبة موزعين على (56) مدرسة كما هو موضح في جدول (1) بحسب إحصائيات قسم التخطيط في مديرية لواء قصبة عمان.

جدول (1)

أعداد طلبة الصف الثامن الأساسي في مديرية لواء قصبة عمان للعام الدراسي 2016/2015

الجنس	عدد المدارس	عدد الطلبة	النسبة
ذكور	23	3483	%47
إناث	33	3921	%53
المجموع	56	7404	%100

عينة الدراسة:

تكوَنت عينة الدراسة من (1015) من الطلبة الذكور والإناث، موزعين على (35) شعبة دراسية في عشر مدارس، تم اختيارهم بالطريقة العشوائية العنقودية حيث كانت وحدة الاختيار هي المدرسة. وقد تمّ حصر المدارس التي يتواجد فيها طلبة الصف الثامن، ومن هذه المدارس تم اختيار أربع مدارس للذكور، وست مدارس للإناث باستخدام الطريقة العشوائية، مع مراعاة التوزيع الجغرافي للمدارس المنتشرة في مواقع مختلفة ومتباعدة من مدينة عمان. ويوضّح الملحق (1) أسماء المدارس والجنس وعدد الصفوف وعدد الطلبة في كل صف (أو شعبة دراسية)، ويوضّح جدول (2) أعداد الطلبة الذين شملتهم عينة الدراسة موزّعين بحسب متغيّرات الجنس.

جدول(2)

توزيع أفرد عينة الدراسة حسب متغيّر الجنس

الجنس	عددالمدارس	عددالطلبة	النسبة
ذكور	4	330	%33
إناث	6	685	%67
المجموع	10	1015	%100

أداة الدراسة:

قام الباحث ببناء اختبار تحصيلي من نوع الاختيار من متعدد من أربعة بدائل في مبحث الرياضيات للصف الثامن الأساسي للفصل الدراسي الأول، والمكوَن من الوحدات الدراسية الثلاث: الأعداد الحقيقية، والمقادير الجبرية، والأنماط والاقترانات. وتم لاحقًا تكوين أربعة اختبارات؛ اثنان منها من نوع الاختيار من أربعة بدائل؛ أحدهما طويل مكوًن من (50) فقرة، والآخر قصير مكوًن من (25) فقرة بناءً على ما توصلت اليه العديد من الدراسات عند مطالعة الأدب النظري، حيث يعود السبب الى أن الاختبارات النفسية والتربوية عادةً ما تستخدم اختبارات بطول (25) فقرة

كاختبارات قصيرة وأخرى بطول (50) فقرة كاختبارات متوسطة (Dawber Roger Carbonaro,2009) المشار اليه في (النصراوين، 2015) في مادة الرياضيات. وكذلك الأخذ بعين الاعتبار خصائص الطلبة النمائية والقدرات العقلية للطلبة في هذا السن حيث اعتبر الباحث أن الاختبار المكون من (50) فقرة طويلاً لهذه الفئة من الطلاب. وبعد ذلك تم استخدام فقرات الاختبارين السابقين نفسها لتكوين اختبارين آخرين من نوع الصواب والخطأ وبنفس العدد من الفقرات؛ أحدهما طويل تكوّن من (50) فقرة، والثاني قصير تكوّن من (25) فقرة. هذا وقد اقتضى بناء فقرات الاختبار ووضعه في الصيغة النهائية الاسترشاد بالأسس العامة المتبعة في بناء الاختبارات التحصيلية التي أوردها (Gronlund & Linn, 1990) كما يأتي:

1) تحديدد الهدف من الاختبار، وهو قياس تحصيل طلبة الصف الثامن الأساسي في مبحث الرياضيات للفصل الدراسي الأول للعام الدراسي 2016/2015.

2) تحليل المحتوى للموضوعات الداخلة في الاختبار، وهي الوحدات الثلاث الأولى من كتاب الرياضيات والتي اشتملت على موضوعات: الأعداد الحقيقية، المقادير الجبرية، الأنماط والاقترانات. ويوضح ملحق (2) تحليل المحتوى لهذه الموضوعات.

3) صياغة النتاجات التعلمية (أهداف المحتوى) موضوع الاختبار كما هو موضّح في ملحق (3).

4) صياغة فقرات اختبار الاختيار من متعدّد (تجمع للفقرات)، مع مراعاة قواعد كتابة الفقرات، حيث اشتملت كل فقرة أربعة بدائل يمثّل إحداها الإجابة الصحيحة، وقد بلغ عدد الفقرات بصورتها الأولية (96) فقرة، وذلك لضمان كفاية عدد الفقرات إذا تم حذف أو دمج الفقرات اعتماداً على نتائج التحكيم، أو المحكات الإحصائية للفقرات، حيث تم وضع عدد من الفقرات لقياس كل هدف من أهداف المحتوى، كما روعي أن لا تعتمد إجابة أية فقرة على إجابة الفقرات الأخرى، وذلك

لتحقيق فرض الاستقلال الموضعي الذي تتطلبه نظرية الاستجابة للفقرة. ويوضّح ملحق (4) الاختبار في صورته الأولية والذي عرض على المحكمين.

5) عرض الصورة الأولية للاختبار مرفقة بتحليل المحتوى والنتاجات التعلمية (أهداف المحتوى) على مجموعة من المحكمين بلغ عددهم (13) محكماً مختصاً (ملحق 5) ذوي خبرة وكفاية في تدريس مبحث الرياضيات، وذلك للتعرف على وجهات نظرهم في الفقرات من حيث وضوح صياغتها اللغوية، وسلامة محتواها العلمي، وقدرتها على قياس الهدف الذي وضعت لقياسه (أي درجة توافق الفقرة مع الهدف الذي تقيسه)، بالإضافة إلى تقييم مستوى الهدف حسب تصنيف بلوم للأهداف المعرفية، وملاءمة بدائل أو مموهات الإجابة لكل فقرة. ويوضح ملحق (6) نموذج التحكيم المرفق مع الاختبار بصورته الأولية، كما يوضّح ملحق (7) نتائج التحكيم وعدد المحكمين المتفقين وغير المتفقين نحو كل فقرة من فقرات الاختبار. وفي ضوء نتائج التحكيم تم حذف الفقرات التي أجمع المحكمون على حذفها، والإبقاء على الفقرات التي اتفق على صحتها عشرة من المحكمين فأكثر من أصل (13 محكم). كما تمّ الأخذ بالاقتراحات والملاحظات التي اقترحها المحكمون، مما ساعد في إعادة النظر ببعض الأهداف، ومراجعة بعض الفقرات وإجراء التعديلات على صياغة الفقرات بهدف إزالة الغموض الذي اعترى صياغة بعض الفقرات، وتصويب الأخطاء اللغوية والإملائية والمطبعية، كما تم اجراء التعديلات على بدائل (مموهات) الفقرات التي أشار إليها بعض المحكمين بأنها ضعيفة. وفي المجمل، تمّ حذف خمس فقرات من الصورة الأولية للاختبار والتي تحمل الأرقام (10، 14، 19، 39، 40) حيث لم يتفق العدد اللازم من المحكمين على ابقائها في الاختبار، وبذلك تكوّنت الصورة الثانية من الاختبار بعد التحكيم من (91) فقرة.

6) الحصول على كتاب لتسهيل مهمة الباحث من خلال مخاطبة إدارة الجامعة الأردنية لوزارة التربية والتعليم من أجل إصدار كتاب تسهيل المهمة (ملحق 8).

7) التجريب الأولي للاختبار: نظراً لطول الاختبار، وحتى لا يشعر الطلبة بالملل والتعب أو اللجوء إلى الإجابة العشوائية دون تفكير أو اهتمام، تمّ تقسيم الفقرات بعد التحكيم وعددها (91 فقرة) إلى اختبارين أحدهما مكوًن من (50) فقرة والآخر مكوًن من (41) فقرة. بعد ذلك تمّ تطبيق هذين الاختبارين على عينة استطلاعية مكوًنة من (163) طالباً وطالبة من مدرسة "محمد إقبال الأساسية للبنين"، ومدرسة "الراشدية الثانوية للبنات"، وهما جزء من مجتمع الدراسة وخارج العينة الرئيسية التي جرى عليها تطبيق الاختبارات بصورتها النهائية. وقد التقى الباحث بنفسه بالطلبة بحضور معلميهم قبل وقت كافٍ من تطبيق الاختبار لتوضيح الهدف من الاختبار وأهميته وضرورة عودة الطلبة إلى مادة الفصل الأول ومراجعتها بشكل جيد، وكان ذلك في منتصف شهر شباط 2016.

8) قام الباحث يوم الاختبار بتوضيح وشرح تعليمات الاختبار للطلبة، وطريقة الإجابة عن الفقرات، وتوضيح كيفية تفريغ الإجابات في النموذج المعد لذلك مع التوضيح بمثال. وتمّ حساب الزمن اللازم للإجابة عن أسئلة الاختبار وذلك بحساب المتوسط الحسابي للزمن الذي استغرقه الطالب الذي أنهى الاختبار أولاً، والزمن الذي احتاجه آخر طالب للانتهاء من الإجابة، وإضافة الزمن اللازم لتوزيع الأوراق على الطلبة، فكان زمن الإجابة للاختبار الأول (50 فقرة) هو (45 دقيقة) وزمن الإجابة للاختبار الثاني (41 فقرة) هو (35 دقيقة)، مع ملاحظة أنه تم تطبيق الاختبارين في يومين متتاليين.

9) تمّ حساب فاعلية البدائل (أو المموهات) لكل فقرة، وذلك من خلال تصحيح أوراق الإجابة وترتيبها تنازليا حسب الدرجة التي حصل عليها الطلبة، ثم تقسيم الطلبة إلى فئتين عليا ودنيا، وكان

عدد كل من الفئتين (44) طالبًا وبنسبة (27%) من الطلبة المتقدمين للاختبار (زيتون، 1996) والبالغ عددهم جميعاً (163) طالبًا وطالبة، وتم حساب فاعلية المموهات (أو البدائل) حسب العلاقة:

فاعلية المموه=عددالذين اختاروا المموه من الفئة العليا ــ عددالذين اختاروا المموه من الفئة الدنيا

عدد الطلبة في إحدى الفئتين (44)

وتجدر الإشارة إلى أن فاعلية الإجابة الصحيحة للفقرة تكون موجبة إذ يفترض أن تجذب طلبة الفئة العليا أكثر من طلبة الفئة الدنيا، بعكس المموه الذي يمثل الإجابة الخطأ فتكون فاعليته سالبة، حيث يفترض أن يجذب طلبة الفئة الدنيا أكثر من طلبة الفئة العليا، والملحق (9) يوضح فاعلية المموهات لفقرات الاختبار المكوّن من (91) فقرة.

10) تمّ رصد إجابات الطلبة بحيث تأخذ الإجابة الصحيحة (1)، وتأخذ الإجابة الخاطئة (0)، والاستعانة بالبرنامج الإحصائي "الرزمة الإحصائية للعلوم الاجتماعية (SPSS)"، وذلك لحساب معامل صعوبة كل فقرة حسب النظرية الكلاسيكية والتي تمثل نسبة الطلبة الذين أجابوا على الفقرة إجابة صحيحة إلى جميع الطلبة الذين حاولوا الإجابة عن الفقرة، وتم أيضاً حساب معامل تمييز كل فقرة، وذلك بحساب معامل الارتباط (بوينت بايسيريال) بين الدرجة على الفقرة والدرجة الكلية على الاختبار للإستعانة بهما في الحكم على فقرات الاختبار، وذلك لجميع أفراد عينة التجريب البالغ عددهم (163) طالبًا وطالبة، كما هو موضّح في الجدول (3).

جدول (3)

معاملات الصعوبة والتمييز لفقرات الاختبار حسب النظرية الكلاسيكية بصورته الثانية (91 فقرة)

الفقرة	معامل الصعوبة	معامل التمييز	الفقرة	معامل الصعوبة	معامل التمييز	الفقرة	معامل الصعوبة	معامل التمييز
1	.82	.41	**33**	.42	**.10**	65	.66	.43
2	.60	.34	34	.59	.56	66	.78	.60
3	.83	.31	35	.45	.54	67	.28	.29
4	.50	.31	**36**	.34	**.09**	68	.45	.29
5	.33	**-.04**	37	.63	.33	69	.30	.39
6	.83	.54	38	.43	.44	**70**	.37	**.07**
7	.78	.33	**39**	**حذفت من قبل المحكمين**		71	.20	.30
8	.43	.49	**40**	**حذفت من قبل المحكمين**		**72**	.35	**.14**
9	.77	.50	41	.75	.43	73	.49	.38
10	**حذفت من قبل المحكمين**		42	.56	.27	74	.42	.29
11	.50	.54	43	.37	.43	75	.57	.41
12	.80	.28	44	.61	.43	76	.63	.42
13	.80	.55	45	.34	.43	77	.38	.40
14	**حذفت من قبل المحكمين**		46	.33	.43	78	.50	.34
15	.25	**-.07**	**47**	.45	**.16**	79	.47	.52
16	.55	.52	**48**	.20	**.07**	80	.55	.39
17	.27	.29	49	.33	.20	81	.79	.50
18	.34	.31	50	.70	.33	82	.46	.27
19	**حذفت من قبل المحكمين**		51	.74	.44	83	.26	.24
20	.70	.46	52	.63	.53	84	.37	.43
21	.84	.39	53	.71	.52	85	.60	.39
22	.33	.41	54	.66	.30	86	.36	.37
23	.39	.32	55	.48	.37	**87**	**.19**	.31
24	.92	.31	56	.50	.30	88	.61	.48
25	.79	.51	57	.57	.28	89	.50	.35
26	.50	.34	58	.28	.19	90	.25	.32
27	.67	.25	59	.72	.44	**91**	.24	**.03**
28	.39	.31	**60**	.55	**.19**	92	.74	.44
29	.52	.41	61	.43	.26	93	.64	.34
30	.43	.47	62	.71	.54	94	.34	.31
31	.28	.47	63	.48	.47	**95**	**.09**	**.10**
32	.43	.30	**64**	**.19**	.40	**96**	**.05**	**-.15**

يُلاحظ من الجدول (3) أن معاملات صعوبة فقرات الاختبار تراوحت بين (0.05 و 0.92) بمتوسط حسابي مقداره (0.50)، بينما تراوحت معاملات التمييز بين (**-0.07 و 0.60**) بمتوسط حسابي مقداره (0.35). وفي ضوء ذلك تم حذف الفقرات التي يقل معامل صعوبتها أو معامل تمييزها عن (0.2) بناءً على ما تم الاطلاع عليه من دراسات في الأدب النظري في هذا المجال. وعليه فقد تم حذف الفقرتين ذات الأرقام (64، 87) لتدني معاملات صعوبتيهما حيث بلغت (0.19) لكل منهما، بالرغم من أن معاملات التمييز لهما مقبولة حيث كانت (0.40، 0.31) على الترتيب، كما تم حذف الفقرات العشر ذات الأرقام (5، 15، 33، 36، 47، 48، 60، 70، 72، 91) وذلك لتدني معاملات تمييزها حيث بلغت (-0.04، -0.07، 0.10، 0.09، 0.16، 0.07، 0.14، 0.03) على الترتيب، بالرغم من أن معاملات الصعوبة لها مقبولة حيث بلغت (0.33، 0.25، 0.42، 0.34، 0.45، 0.2، 0.55، 0.37، 0.35، 0.24) على الترتيب، وكذلك تم حذف الفقرتين (95، 96) لتدني معاملات صعوبتهما ومعاملات تمييزهما معًا حيث بلغت معاملات صعوبتهما (0.09، 0.05) على الترتيب، وبلغت معاملات تمييزهما (0.10، -0.15) على الترتيب. وعليه، بلغ مجموع عدد الفقرات التي تم حذفها سواء بسبب تدني معامل صعوبتها أو تدني معامل تمييزها أو تدني المعاملين معاً (14) فقرة، ليصبح عدد فقرات الاختبار بصورته الثالثة والتي تمّ الاحتفاظ بها (77) فقرة. وفي المرحلة الرابعة، قام الباحث بفرز الفقرات التي تمّ الاحتفاظ بها بهدف تشكيل صورتين من الاختبار مجموع فقراتهما (75 فقرة)؛ هما صورة الاختبار الطويل الذي تكون من (50) فقرة، وصورة الاختبار القصير الذي تكوّن من (25) فقرة وبذلك يبقى فقرتان زائدتان عن الحاجة هما (29، 38). ويوضّح ملحق (10) أرقام الفقرات في الاختبار بصورته الرابعة وأرقامها الجديدة في الاختبارين (25 فقرة، 50 فقرة). وتجدر الإشارة إلى أنه روعي في توزيع الفقرات على صورتي الاختبار ربط مستويات النتاجات التعليمية أو

الأهداف المصنّفة وفق تصنيف بلوم للأهداف المعرفية بمحتوى المادة الدراسية موضوع الاختبار أو ما يطلق عليه جدول مواصفا الاختبار كما هو موضّح في ملحق (11).

الخصائص السيكومترية لأداة الدراسة:

صدق الاختبار:

تم التحقق من صدق نتائج الاختبار انطلاقًا مما حددته رابطة علم النفس الأمريكية (American Psychological Association, 1974) من وجود ثلاثة مصطلحات رئيسة يشير كل منها إلى نوع من أنواع الصدق وهي (عودة،2004):

1- صدق المحتوى بنوعيه الظاهري والعيني.

2- الصدق المرتبط بمحك من خلال الصدق التلازمي.

3- صدق البناء من خلال التحليل العاملي.

وفيما يأتي تفصيل للإجراءات المتبعة للتوصّل إلى مؤشرات حول صدق نتائج الاختبار الآنفة الذكر، علمًا بأنه تم تطبيق هذه الإجراءات على كل من الاختبارين الطويل (50 فقرة) والقصير (25 فقرة) من نوع الاختيار من متعدد.

1- صدق المحتوى:

أ- الصدق الظاهري Face Validity: يُمكن القول أنه تحقّق للاختبار بصورتيه؛ الطويل والقصير دلالة صدق المحتوى (الصدق الظاهري) من خلال عرض الاختبار على ثلاثة عشر من المحكمين المختصين. وبناء عليه تم اختيار فقرات الاختبار التي توافق عليها عشرة فأكثر من المحكمين من

حيث سلامة الفقرة لغويًا وعلميًا، ومناسبتها لقياس سمة التحصيل المطلوبة للاختبار، حيث تم التبصر في مضمون كل فقرة من فقرات الاختبار والحكم على مدى علاقتها بالمادة الدراسية المعنية.

ب- الصدق العيني Sampling Validity: يُمكن القول بأنه تحقق للاختبار دلالة صدق المحتوى (الصدق العيني) للاختبار من خلال تحديد المجال والموضوعات الدراسية التي يغطيها الاختبار بشكل دقيق وواضح، وإعداد جدول مواصفات الاختبار، حيث تمّ حصر الموضوعات وتحديد الأهمية النسبية لكل منها وتمثيلها في الاختبار بما يتناسب مع أهميتها ووزنها.

2- الصدق المرتبط بالمحك: تم التحقّق من الصدق المرتبط بالمحك باستخدام الصدق التلازمي Concurrent Validity حيث تم حساب معامل الارتباط بين درجات كل من الاختبارين الطويل (50 فقرة)، والقصير (25 فقرة) مع درجات الطلبة المدرسية في مادة الرياضيات في الفصل الدراسي الأول المعتمدة من معلم المادة كمحك، وقد بلغت قيمة معامل الارتباط بين الدرجات على الاختبار الطويل (50 فقرة) ودرجات الطلبة المدرسية (0.843)، ومعامل الارتباط بين الدرجات على الاختبار القصير(25 فقرة) ودرجات الطلاب المدرسية (0.762)، وهذا يوضح أن قيم معاملات الارتباط موجبة ومرتفعة أي أن صورتي الاختبار تتمتع بدلالة صدق محكي مرتفع.

3- صدق البناء: ويهدف هذا النوع من الصدق إلى التحقق من أن الاختبار يقيس السمة التي أعد لقياسها، والسمة المقاسة هنا هي التحصيل في مادة الرياضيات، وفي هذا الاختبار تم التحقق من الصدق العاملي للاختبارات باستخدام التحليل العاملي Factor Analysis باستخدام طريقة المكونات الأساسية Principal Component Analysis كواحد من المؤشرات الإحصائية المستخدمة في الكشف عن صدق البناء، حيث تم إجراء التحليل العاملي

لاستجابات الأفراد على الاختبارين باستخدام برنامج الرزمة الإحصائية للعلوم الاجتماعية (SPSS) وذلك للتعرف على العوامل المسؤولة عن الأداء في كلا الاختبارين، والتأكد من أن الأداء يعود لعامل القدرة الرياضية (التحصيل). ويبين الجدولان (4 ، 5) قيم الجذور الكامنة Eigenvalue، ونسبة التباين المفسر Explained Variance، ونسبة التباين التراكمي للعوامل التي لها جذر كامن أكبر من الواحد صحيح، وجرى تدوير العوامل باستخدام طريقة التدوير المتعامد Varimax Rotation للعوامل التي كانت قيمة الجذر الكامن لها أكبر من واحد صحيح لتحسين تفسير العوامل الناتجة من التحليل قبل التدوير.

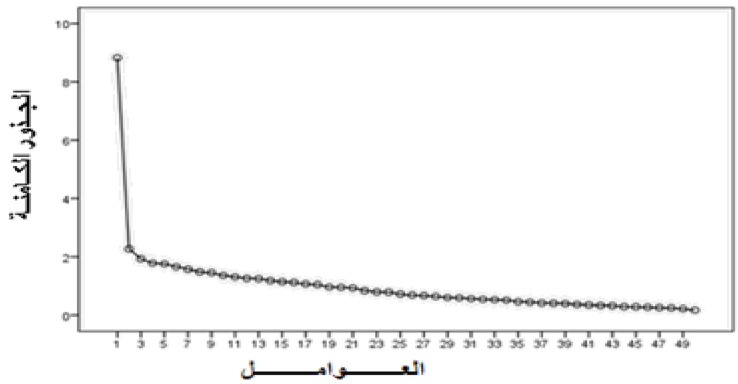
جدول (4)

قيم جميع الجذور الكامنة الأكبر من الواحد صحيح ونسبة التباين المفسر للعوامل للاختبار الطويل (50 فقرة)

العامل	الجذر الكامن	نسبة التباين المفسر %	نسبة التباين المفسر التراكمي %
1	8.826	17.651	17.651
2	2.268	4.536	22.187
3	1.933	3.866	26.052
4	1.785	3.571	29.623
5	1.759	3.518	33.141
6	1.659	3.319	36.460
7	1.575	3.149	39.609
8	1.481	2.962	42.571
9	1.452	2.903	45.474
10	1.363	2.725	48.199

يتّضح من جدول (4) أعلاه وجود عشرة عوامل لها جذور كامنة قيمها تزيد عن الواحد الصحيح، وفسرت ما يقارب (48%) من التباين الكلي في الاختبار الطويل (50 فقرة)،

وعند قسمة الجذر الكامن للعامل الأول على الجذر الكامن للعامل الثاني يكون خارج القسمة يساوي (3.90)، وهي نسبة تزيد عن ضعفين، ومن ناحية ثانية كانت نسبة الفرق بين الجذرين الكامنين الأول والثاني إلى الفرق بين الجذرين الكامنين الثاني والثالث مرتفعة حيث بلغت هذه النسبة (19.6)، وتعد هذه القيم كما يشير هاتي (Hattie, 1985) مؤشراً على أحادية البعد، وهذا يعني أن الاختبار يقيس سمة واحدة وهي التحصيل في الرياضيات، مما يسمح باستخدام نماذج الاستجابة للفقرة في تقدير معالم الفقرات وقدرة الأفراد. كما يبين الشكلان (1،2) الرسوم البيانية التي تمثل الجذور الكامنة للعوامل جميعها لبيانات الاختبارين، وهو ما يعرف باختبار فرز العوامل (Scree plot).



شكل (1): الرسم البياني للعلاقة بين العوامل والجذور الكامنة للاختبار الطويل (50 فقرة).

يتّضح من الشكل (1) الذي يمثل الجذور الكامنة للعوامل جميعها بيانياً لبيانات الاختبار الطويل (50 فقرة) –عند تطبيقه على عينة التجريب– أن مقدار الميل بين العامل الأول والعامل الثاني كان كبيراً جداً مقارنة بالميل بين العامل الثاني والعامل الثالث، أي أن المنحنى ينحدر بشدة بعد العامل الأول، ثم يبدأ ميل الخط المنحني بالتغيّر بين النقطتين المناظرتين للعاملين الثاني والثالث، وقد أصبحت قيم الجذور الكامنة للعوامل بدءاً من العامل الثالث متقاربة، وهذا بمثابة دليل

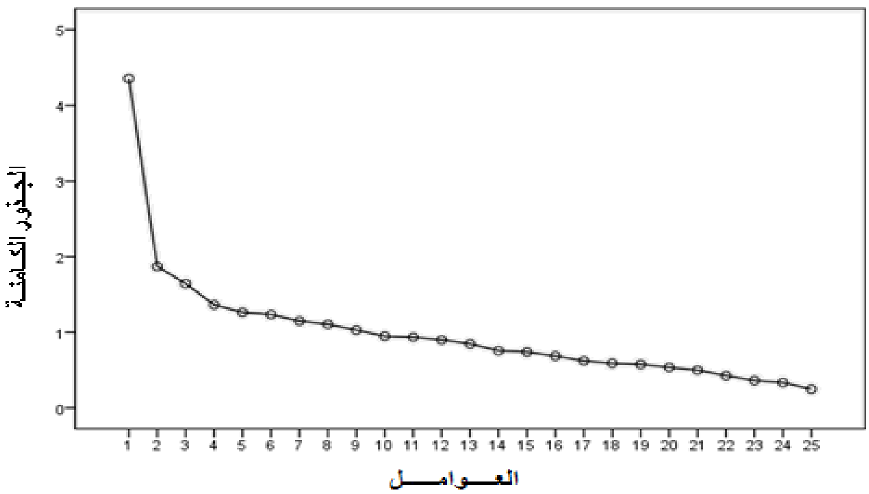
قوي على وجود عامل مسيطر أو سائد على بقية العوامل الأخرى وهو عامل التحصيل في الرياضيات.

جدول (5)

قيم جميع الجذور الكامنة الأكبر من الواحد صحيح ونسبة التباين المفسر للعوامل للاختبار القصير (25 فقرة)

العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي
1	4.353	17.414	17.414
2	1.869	7.477	24.891
3	1.643	6.571	31.462
4	1.363	5.450	36.912
5	1.262	5.049	41.961
6	1.235	4.939	46.900
7	1.148	4.593	51.493
8	1.106	4.424	55.917

يتّضح من جدول (5) وجود ثمانية عوامل لها جذور كامنة أكبر من واحد صحيح، وفسرت ما يقارب (56%) من التباين الكلي في الاختبار القصير (25 فقرة)، وعند قسمة الجذر الكامن للعامل الأول على الجذر الكامن للعامل الثاني يكون خارج القسمة يساوي (2.33)، وهي نسبة تزيد عن ضعفين، ومن ناحية ثانية كانت نسبة الفرق بين الجذرين الكامنين الأول والثاني إلى الفرق بين الجذرين الكامنين الثاني والثالث مرتفعة حيث بلغت هذه النسبة (11)، وتعد هذه القيم كما يشير هاتي (Hattie,1985) مؤشراً على أحادية البُعد، وهذا يعني أن الاختبار يقيس سمة واحدة وهي التحصيل في الرياضيات، مما يسمح باستخدام نماذج الاستجابة للفقرة في تقدير معالم الفقرات وقدرة الأفراد.



شكل (2): الرسم البياني للعلاقة بين العوامل والجذور الكامنة للاختبار القصير (25 فقرة).

يتّضح من شكل (2) الذي يمثل الجذور الكامنة للعوامل جميعها بيانياً لبيانات الاختبار القصير (25 فقرة) أن مقدار الميل بين العامل الأول والعامل الثاني كان كبيراً جداً مقارنة بالميل بين العامل الثاني والعامل الثالث، أي أن المنحنى ينحدر بشدة بعد العامل الأول، ثم يبدأ ميل الخط المنحني بالتغيّر بين النقطتين المناظرتين للعاملين الثاني والثالث، وقد أصبحت قيم الجذور الكامنة للعوامل بدءاً من العامل الثالث متقاربة، وهذا بمثابة دليل قوي على وجود عامل مسيطر أو سائد على بقية العوامل الأخرى وهو عامل التحصيل في الرياضيات.

ولدى حساب معامل الارتباط بين الفقرة والعامل الأول السائد (تشبع الفقرات على العامل السائد) لكل اختبار من الاختبارين الطويل (50 فقرة)، والقصير (25 فقرة) كما هو موضّح في الجدول (6) الآتي تبيّن أن هذه التشبعات تراوحت بين (0.290 ـ 0.583) للاختبار الطويل، في حين تراوحت للاختبار القصير بين (0.291 ــ 0.709)، ويلاحظ أن الحد الأدنى لهذه التشبعات بلغ (0.30) تقريباً، وهو الحد الأدنى المقبول للتشبع حتى يكون العامل سائداً (Harlow,2005)، وهذا دليل آخر على وجود العامل السائد وهو التحصيل في الرياضيات.

جدول (6)

تشبعات الفقرات على العامل الأول (السائد) للاختبارين الطويل والقصير

الاختبار الطويل (50 فقرة)				الاختبار القصير (25 فقرة)	
الفقرة		**الفقرة**		**الفقرة**	
1	0.424	**26**	0.290	**1**	0.414
2	0.292	**27**	0.367	**2**	0.353
3	0.572	**28**	0.439	**3**	0.601
4	0.348	**29**	0.316	**4**	0.622
5	0.297	**30**	0.460	**5**	0.392
6	0.488	**31**	0.294	**6**	0.296
7	0.404	**32**	0.293	**7**	0.359
8	0.398	**33**	0.317	**8**	0.317
9	0.342	**34**	0.470	**9**	0.548
10	0.558	**35**	0.577	**10**	0.322
11	0.485	**36**	0.375	**11**	0.300
12	0.470	**37**	0.290	**12**	0.291
13	0.583	**38**	0.295	**13**	0.334
14	0.558	**39**	0.419	**14**	0.297
15	0.351	**40**	0.385	**15**	0.642
16	0.473	**41**	0.317	**16**	0.490
17	0.434	**42**	0.529	**17**	0.709
18	0.437	**43**	0.365	**18**	0.405
19	0.428	**44**	0.545	**19**	0.478
20	0.464	**45**	0.414	**20**	0.296
21	0.529	**46**	0.385	**21**	0.320
22	0.320	**47**	0.507	**22**	0.385
23	0.455	**48**	0.325	**23**	0.485
24	0.554	**49**	0.349	**24**	0.298
25	0.551	**50**	0.325	**25**	0.296

ثبات الاختبار:

بهدف التحقق من ثبات كل اختبار من الاختبارين؛ الطويل المكوَن من (50 فقرة)، والقصير المكوّن من (25 فقرة) من نوع الاختيار من متعدد، تم حساب الثبات بطريقة الاتساق الداخلي، باستخدام معادلة كودر – ريتشاردسون (KR–20) التي تستخدم في حالة الاختبارات من نوع الإجابة المنتقاة (الموضوعية). وقد بلغ معامل الاتساق الداخلي للاختبار الطويل 0.897، بينما بلغ معامل الاتساق الداخلي للاختبار القصير 0.773، وهي معاملات ثبات مقبولة، كونها تزيد عن (0.70) كما يشير كل من دينيك وتفكول (Dennick & Tavakol, 2011)، إذ أن قيم معامل الثبات التي تقل عن (0.70) تشير إلى ارتباطات ضعيفة بين فقرات الاختبار وبالتالي عدم تجانسها. من ناحية ثانية، يُعد معامل الثبات المرتفع مؤشرًا على تحقق افتراض أحادية البعد في الاختبار.

التحقق من افتراضات نظرية الاستجابة للفقرة:

– التحقق من افتراضات أحادية البُعد والاستقلال الموضعي:

من خلال العودة إلى التحليل العاملي السابق والجدولين (4، 5) السابقين لكل اختبار من الاختبارين (50 فقرة، 25 فقرة)، والتحقق من وجود عامل سائد ومسيطر، وتحقق النسب اللازمة لتحقيق افتراض أحادية البُعد وذلك من خلال الشكلين (1، 2)، وكذلك عند حساب معامل ثبات الاتساق الداخلي والذي تشير نسبه المحسوبة إلى تحقق افتراض أحادية البُعد، كل ذلك يُعد مؤشرات على أن الاختبارين يقيسان سمة واحدة وهي سمة التحصيل في الرياضيات.

أما افتراض الاستقلال الموضعي، فيُعد مكافئاً لافتراض أحادية البُعد وفق ما أشار اليه هامبلتون وسوامينثان (Hambleton & Swaminathan, 1985)، حيث عدَا هذين الافتراضين متلازمين، أي إذا تحقق الافتراض الأول فإن الافتراض الثاني يتحقق أيضاً.

- **التحقق من افتراض التحرر من السرعة:**

بالنسبة لافتراض التحرر من السرعة في الأداء، فقد أنهى جميع الطلبة الإجابة على الاختبارين ضمن الوقت المحدد للإجابة، ولم يحتج أي من الطلبة وقتاً اضافياً لإكمال إجابته، وهذا يعني أن إخفاق الطلبة في الإجابة على فقرات الاختبارين عائد إلى انخفاض قدراتهم وليس بسبب عامل السرعة، أي ليس بسبب ضيق وقت الاختبار.

تشكيل الصور المختلفة من الاختبار:

بعد القيام بالإجراءات السابقة على الاختبارين من نوع الاختيار من متعدد ذات البدائل الأربعة أحدهما طويل مكوَن من (50 فقرة)، والآخر قصير مكوَن من (25 فقرة)، قام الباحث بإعادة صياغة نفس الفقرات في الاختبارين لتصبح الإجابة على الفقرة من نوع الصواب والخطأ، ولم يجر على الاختبارين الجديدين (من نوع الصواب والخطأ) الإجراءات السابقة نفسها، بل تم الاكتفاء بالإجراءات التي تمت على اختباريَ الاختيار من متعدد، وذلك حتى لا يضطر الباحث إلى حذف أيّ من الفقرات والحفاظ على العدد نفسه من الفقرات. وفي الصورة النهائية أصبح لدى الباحث أربعة اختبارات هي:

أ- اختبار طويل مكوَن من (50 فقرة) من نوع الاختيار من أربعة بدائل، موضح في ملحق (12).

ب- اختبار طويل مكوَن من (50 فقرة) من نوع الصواب والخطأ، موضح في ملحق (13).

جـ- اختبار قصير مكوَن من (25 فقرة) من نوع الاختيار من متعدد، موضح في ملحق (14).

د- اختبار قصير مكوَن من (25 فقرة) من نوع الصواب والخطأ، موضح في ملحق (15).

وبالنسبة للتصحيح، فقد استخدم الباحث مفتاح الإجابة الموضح في ملحق (16) للاختبارات الأربعة لدى تصحيح إجابات الطلبة. أما بالنسبة لحجم العينة، فقد تم تقسيم حجم عينة الدراسة إلى ثلاثة أقسام (أو أحجام) هي:

أ- عينة صغيرة حجمها (250) مفحوص.

ب- عينة متوسطة حجمها (500) مفحوص.

جـ- عينة كبيرة حجمها (1015) مفحوص.

الإجراءات على عينة الدراسة:

تمّ تطبيق الاختبارات الأربعة المشار اليها وبصورها النهائية على عينة التطبيق الرئيسية الواردة في جدول (2)، وذلك بهدف الوصول إلى نتائج الدراسة، وحسب الخطوات الآتية:

ـ بعد الحصول على الموافقة الرسمية والمشار اليها سابقًا من خلال كتاب تسهيل مهمة في ملحق (8) قام الباحث بزيارة إلى كل مدرسة من المدارس التي وقع عليها الاختيار كمدارس مشاركة في الدراسة الواردة في ملحق (1)، والالتقاء بطلبتها قبل وقت كافٍ من موعد غقد الاختبار وضرورة الاستعداد الجيد وبذل أقصى الجهد، والتنسيق مع إدارة كل مدرسة ومعلمي ومعلمات مادة الرياضيات للصف الثامن الأساسي في تلك المدارس.

ـ شرح طريقة الإجابة وتوضيح تعليمات الاختبار للطلبة، ثم توزيع أوراق الاختبار على الطلبة مرفقاً

بنموذج الإجابة لكل اختبار، وتم عقد الاختبارات الأربعة على يومين متتاليين، حيث تقدم الطلبة في اليوم الأول للاختبارين؛ الاختيار من متعدد المكوَن من (50 فقرة)، ويليه في الجلسة الثانية اختبار الصواب والخطأ المكوَن من (25 فقرة). وفي اليوم التالي تقدم الطلبة لاختبار الصواب والخطأ المكوَن من (50 فقرة) ويليه في الجلسة الثانية تقدم الطلبة لاختبار الاختيار من متعدد المكوَن من (25 فقرة)، وتم ذلك تحت إشراف الباحث وبالتعاون مع معلمي ومعلمات المادة، حيث استمر تطبيق هذه الاختبارات عشرين يومًا.

ـ جمع الأوراق من الطلبة واستثناء الأوراق غير مكتملة الإجابة، حيث كانت أعداد الطلبة الموضحة في جدول (2) هي الأعداد النهائية، بعد استثناء (30) ورقة غير مكتملة الإجابة من جميع الصفوف وللاختبارات الأربعة.

ـ تصحيح الأوراق للاختبارات الأربعة يدويًا، بوضع علامة (1) للإجابة الصحيحة على الفقرة وعلامة (0) للإجابة الخاطئة، ذلك وفق نموذج تصحيح أعده الباحث.

ـ فرز أوراق الإجابة حسب اسم الطالب (أي تجميع أوراق إجابة كل طالب) حيث تكوّن في النهاية أربع علامات لكل طالب بعدد الاختبارات التي تقدم لها ليُصار بعد ذلك إلى إعطاء الطلبة أرقام متسلسلة من 1 إلى 1015، ثم تم إدخال درجات الطلاب إلى برنامج (Ecxel) حسب أرقامهم المتسلسلة حتى يتسنى نقلها إلى برامج التحليل الإحصائي الأخرى المستخدمة في هذه الدراسة وهي (Bilog-Mg,SPSS) بهدف تحليلها.

ـ الاستعانة ببرنامج الرزمة الإحصائية للعلوم الاجتماعية (SPSS) في الاختيار العشوائي للعينات، حيث تم اختيار عينة عشوائية صغيرة مكوَنة من (250) مفحوصًا، وعينة متوسطة مكوَنة من

(500) مفحوص، والعينة الكبيرة وهي عينة الدراسة جميعها والمكونة من (1015) مفحوصًا.

ـ إجراء التحليل العاملي لنتائج الاختبارات الأربعة للتحقق من افتراض أحادية البعد، وافتراض الاستقلال الموضعي، وذلك من خلال الاستعانة ببرنامج الرزمة الإحصائية للعلوم الاجتماعية (SPSS)، حيث تم التحقق من هذين الافتراضين وعلى مستويات حجوم العينة الثلاثة ، ويلي ذلك التوضيح من خلال الرسوم البيانية لاختبار فرز العوامل (Scree Plot)، كما هو موضّح في الجداول (7، 8، 9) جميع العوامل التي قيم جذورها الكامنة أكبر من الواحد صحيح طبقاً لمحك كايزر (Kaiser) المرتبط بنموذج المكونات الأساسية، بالإضافة إلى نسبة التباين المفسر لكل منها وذلك لكلا النوعين من الاختبارات (الاختيار من متعدد والصواب والخطأ) بطولين مختلفين (25 فقرة، 50 فقرة)، علمًا بأنه جرى تدوير للعوامل بطريقة التدوير المتعامد (Varimax Rotation)

في جدول (7) الآتي من المُلاحظ أن قيمة الجذر الكامن للعامل الأول بلغت للاختبار القصير من نوع اختيار من متعدد (5.133) ويفسر ما نسبته 20.532% من التباين الكلي من اجابات الطلبة، وهي قيمة مرتفعة إذا ما قورنت مع قيم الجذور الكامنة لبقية العوامل، أما قيمة الجذر الكامن للعامل الثاني بلغت 1.496 ويفسر ما نسبته 5.984% من التباين الكلي، بمعنى أن العامل الأول فسر ما يزيد عن ضعفيَ ما فسره العامل الثاني، ويُلاحظ أيضاً أن نسبة التباين المفسر لكل العوامل المتبقية متقاربة، بمعنى أنه يوجد شبه استقرار في نسب التباين المفسر لجميع العوامل باستثناء العامل الأول، وهذا مؤشر على تحقق افتراض احادية البعد للمقياس (Hattie,1985)، أي أن الاختبار يقيس سمة واحدة وهي التحصيل في الرياضيات. ويمكن تعميم ما سبق على باقي الاختبارات بالنسبة للعينة المكونة من (250) مفحوصاً، ويُلاحظ أن عدد الجذور الكامنة يتناسب تناسباً طردياً مع عدد فقرات الاختبار.

جدول (7)

قيم جميع العوامل وجذورها الكامنة – الأكبر من الواحد صحيح – ونسبة التباين المفسر لكل منها للاختبارات الأربعة لعينة مفحوصين حجمها (250) مفحوصًا

	نوع الفقرات							
	الاختيار من متعدد				الصواب والخطأ			
طول الاختبار	25 فقرة		50 فقرة		25 فقرة		50 فقرة	
العامل	الجذر الكامن	نسبة التباين المفسر	الجذر الكامن	نسبة التباين المفسر	الجذر الكامن	نسبة التباين المفسر	الجذر الكامن	نسبة التباين المفسر
1	5.133	20.532	5.330	10.660	4.007	16.029	8.120	16.241
2	1.496	5.984	2.409	4.818	1.895	7.579	2.991	5.983
3	1.317	5.269	2.266	4.531	1.648	6.593	2.006	4.012
4	1.259	5.036	2.073	4.145	1.373	5.493	1.784	3.567
5	1.207	4.827	2.026	4.052	1.311	5.242	1.638	3.276
6	1.190	4.761	1.845	3.691	1.221	4.885	1.536	3.072
7	1.103	4.411	1.787	3.574	1.186	4.743	1.445	2.894
8	1.030	4.120	1.706	3.412	1.100	4.400	1.413	2.826
9			1.640	3.280	1.056	4.223	1.363	2.727
10			1.575	3.149			1.305	2.609
11			1.463	2.926			1.216	2.431
12			1.447	2.894			1.152	2.303
13			1.363	2.725			1.097	2.195
14			1.285	2.570			1.091	2.182
15			1.237	2.474			1.055	2.109
16			1.123	2.246			1.040	2.081
17			1.083	2.165				
18			1.058	2.115				
19			1.001	2.002				

جدول (8)

قيم جميع العوامل وجذورها الكامنة – الأكبر من الواحد صحيح – ونسبة التباين المفسر لكل منها للإختبارات الأربعة لعينة مفحوصين حجمها (500) مفحوصًا

	نوع الفقرات							
	الاختيار من متعدد				الصواب والخطأ			
طول الاختبار	25 فقرة		50 فقرة		25 فقرة		50 فقرة	
العامل	الجذر الكامن	نسبة التباين المفسر%	الجذر الكامن	نسبة التباين المفسر%	الجذر الكامن	نسبة التباين المفسر	الجذر الكامن	نسبة التباين المفسر
1	6.128	24.512	7.537	15.074	6.739	26.957	4.301	8.602
2	1.392	5.570	1.868	3.736	1.454	5.816	1.856	3.713
3	1.210	4.838	1.758	3.515	1.379	5.517	1.737	3.474
4	1.132	4.528	1.488	2.976	1.143	4.572	1.637	3.275
5	1.047	4.189	1.445	2.890	1.124	4.497	1.563	3.126
6	1.030	4.120	1.422	2.844	1.063	4.253	1.544	3.088
7			1.340	2.680	1.038	4.151	1.499	2.997
8			1.298	2.597			1.473	2.947
9			1.233	2.466			1.391	2.782
10			1.220	2.441			1.355	2.709
11			1.188	2.376			1.293	2.586
12			1.133	2.266			1.250	2.499
13			1.111	2.222			1.221	2.441
14			1.069	2.137			1.182	2.363
15			1.053	2.105			1.179	2.359
16			1.031	2.061			1.146	2.292
17			1.005	2.010			1.114	2.228
18							1.069	2.137
19							1.037	2.073

في جدول (8) أعلاه من المُلاحظ أن قيمة الجذر الكامن للعامل الأول بلغت للاختبار القصير من نوع اختيار من متعدد (6.128) ويفسر ما نسبته 24.512% من التباين الكلي من اجابات الطلبة، وهي قيمة مرتفعة إذا ما قورنت مع قيم الجذور الكامنة لبقية العوامل، أما قيمة الجذر الكامن للعامل الثاني بلغت 1.392 ويفسر ما نسبته 5.570% من التباين الكلي، بمعنى أن العامل الأول فسر ما يزيد عن ضعفيَ ما فسره العامل الثاني، ويُلاحظ أيضاً أن نسبة التباين المفسر لكل العوامل المتبقية متقاربة، بمعنى أنه يوجد شبه استقرار في نسب التباين المفسر لجميع العوامل باستثناء العامل الأول، وهذا مؤشر على تحقق افتراض احادية البعد للمقياس (Hattie,1985)، أي أن الاختبار يقيس سمة واحدة وهي التحصيل في الرياضيات. ويمكن تعميم ما سبق على باقي الاختبارات بالنسبة للعينة المكونة من (500) مفحوصاً، ومن الملاحظ أيضاً أن عدد الجذور الكامنة يتناسب تناسباً طردياً مع عدد فقرات الاختبار.

في جدول (9) الآتي من المُلاحظ أن قيمة الجذر الكامن للعامل الأول بلغت للاختبار القصير من نوع اختيار من متعدد (4.899) ويفسر ما نسبته 19.595% من التباين الكلي من اجابات الطلبة، وهي قيمة مرتفعة إذا ما قورنت مع قيم الجذور الكامنة لبقية العوامل، أما قيمة الجذر الكامن للعامل الثاني بلغت (1.396) ويفسر ما نسبته 5.583% من التباين الكلي، بمعنى أن العامل الأول فسر ما يزيد عن ضعفيَ ما فسره العامل الثاني، ويُلاحظ أيضاً أن نسبة التباين المفسر لكل العوامل المتبقية متقاربة، بمعنى أنه يوجد شبه استقرار في نسب التباين المفسر لجميع العوامل باستثناء العامل الأول، وهذا مؤشر على تحقق افتراض احادية البعد للمقياس (Hattie,1985)، أي أن الاختبار يقيس سمة واحدة وهي التحصيل في الرياضيات. ويمكن تعميم ما سبق على باقي الاختبارات بالنسبة للعينة المكونة من (1015) مفحوصاً.

جدول (9)

قيم جميع العوامل وجذورها الكامنة – الأكبر من الواحد صحيح – ونسبة التباين المفسر لكل منها للإختبارات الأربعة لعينة مفحوصين حجمها (1015 مفحوص)

	نوع الفقرات							
	الاختيار من متعدد				الصواب والخطأ			
طول الاختبار	25 فقرة		50 فقرة		25 فقرة		50 فقرة	
العامل	الجذر الكامن	نسبة التباين المفسر%	الجذر الكامن	نسبة التباين المفسر%	الجذر الكامن	نسبة التباين المفسر	الجذر الكامن	نسبة التباين المفسر
1	4.899	19.595	7.112	14.223	4.730	18.921	4.137	8.274
2	1.396	5.583	1.714	3.427	1.578	6.313	1.708	3.417
3	1.217	4.869	1.542	3.085	1.530	6.119	1.574	3.147
4	1.106	4.425	1.459	2.917	1.233	4.931	1.521	3.042
5	1.083	4.332	1.375	2.750	1.196	4.784	1.483	2.967
6	1.017	4.069	1.300	2.600	1.176	4.705	1.391	2.782
7			1.267	2.534	1.116	4.463	1.366	2.732
8			1.230	2.460	1.065	4.259	1.340	2.679
9			1.199	2.98			1.270	2.541
10			1.151	2.301			1.254	2.507
11			1.131	2.262			1.238	2.476
12			1.097	2.194			1.212	2.424
13			1.082	2.163			1.192	2.384
14			1.053	2.106			1.167	2.333
15			1.032	2.063			1.124	2.247
16							1.109	2.218
17							1.083	2.167
18							1.066	2.132

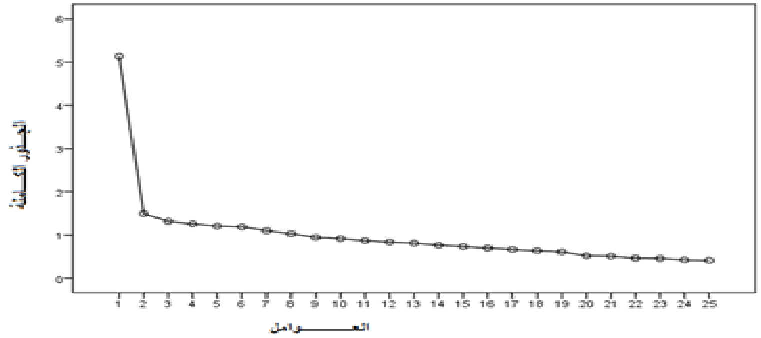
عند التمعن في الجدول (10) الآتي الذي يلخّص النتائج السابقة في الجداول (7، 8، 9) من خلال قسمة الجذر الكامن الأول على الجذر الكامن الثاني في جميع الحالات يتبين أن النسبة تزيد عن الضعفين، وكذلك فإن نسبة الفرق بين الجذرين الكامنين الأول والثاني إلى الفرق بين الجذرين الكامنين الثاني والثالث (نسبة الفروق) هي نسب مرتفعة في جميع الحالات، وتعد هذه النسب المرتفعة مؤشرًا على أحادية البعد؛ أي أن الاختبارات الأربعة وباختلاف عدد فقراتها ونوع هذه الفقرات تقيس سمة واحدة في ظل اختلاف عدد المفحوصين وهذه السمة هي سمة التحصيل .

جدول (10)

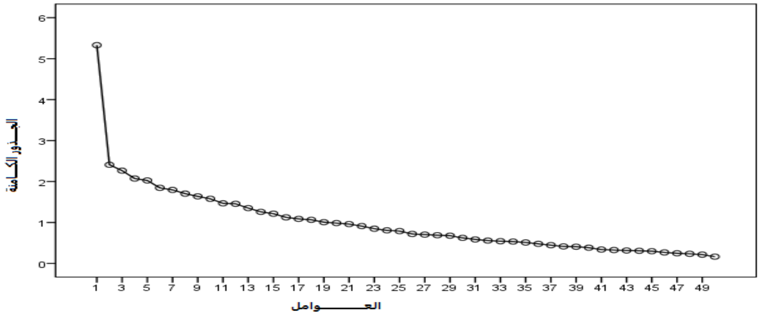
نسبة الجذر الكامن الأول إلى الجذر الكامن الثاني، ونسبة الفرق بين الجذرين الأول والثاني إلى الفرق بين الجذرين الثاني والثالث للاختبارات الأربعة (نسبة الفروق) للعينات (250،1015،500)

	نوع الفقرات							
	الاختيار من متعدد				الصواب والخطأ			
طول الاختبار	25 فقرة		50 فقرة		25 فقرة		50 فقرة	
حجم العينة	**نسبة الجذر الأول إلى الثاني**	**نسبة الفروق**	**نسبة الجذر الأول إلى الثاني**	**نسبة الفروق**	**نسبة الجذر الأول إلى الثاني**	**نسبة الفروق**	**نسبة الجذر الأول إلى الثاني**	**نسبة الفروق**
250	3.43	20.32	2.21	20.43	2.11	8.55	2.72	5.2
500	4.4	26.02	4.03	51.54	4.63	70.47	2.32	20.55
1015	3.51	19.57	4.15	31.38	3	65.67	2.42	18.13

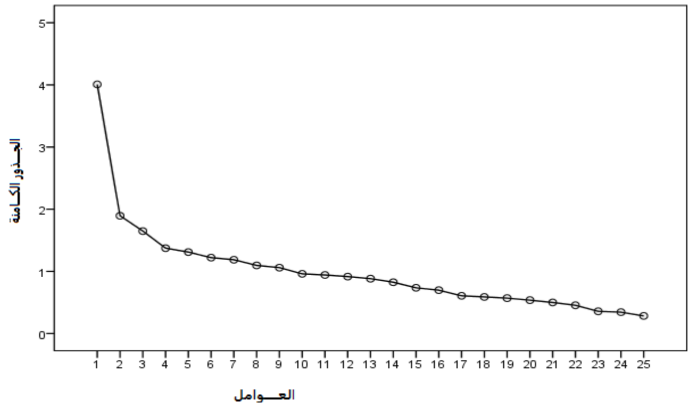
ولتعزيز هذه النتيجة، وباستخدام التحليل العاملي من خلال الرزمة الإحصائية للعلوم الاجتماعية (SPSS)، تم تمثيل العلاقات البيانية بين العوامل والجذور الكامنة وهو ما يعرف باختبار فرز العوامل (Scree Plot)، وذلك للاختبارات الأربعة في ضوء عيّنات بحجوم مختلفة كما في الأشكال (3، 4، 5) الآتية:



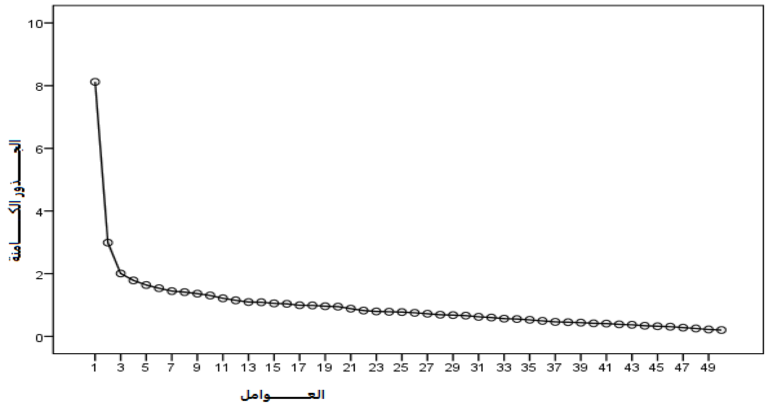
الاختيار من متعدد (25 فقرة)



الاختيار من متعدد (50 فقرة)

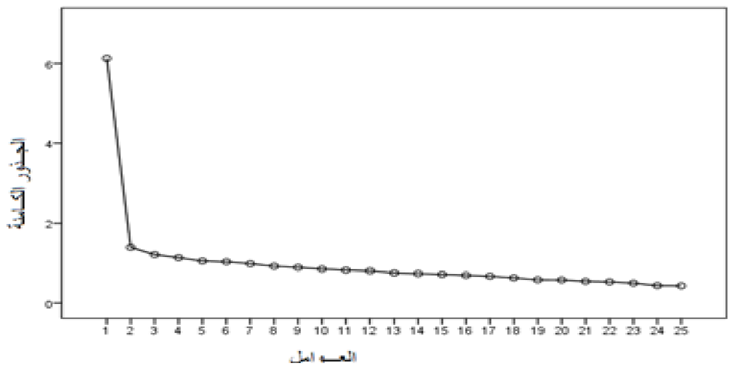


صواب وخطأ (25 فقرة)

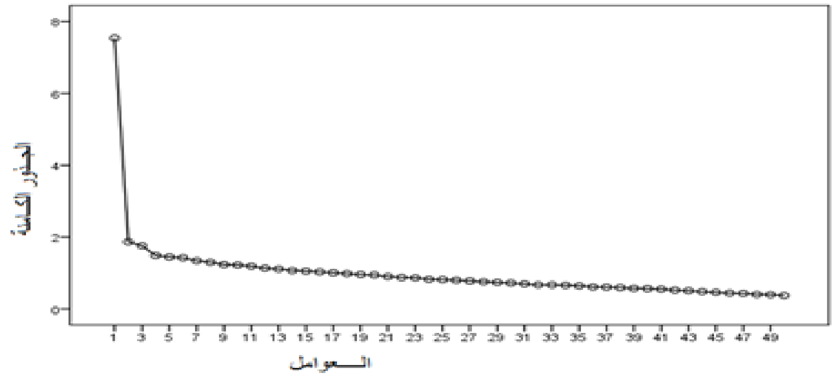


صواب وخطأ (50 فقرة)

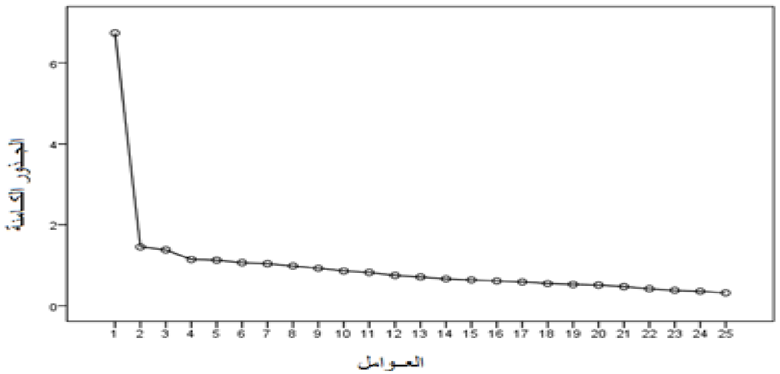
شكل (3): الرسوم البيانية للعلاقة بين قيم الجذور الكامنة والعوامل للإختبارات الأربعة لعينة من المفحوصين حجمها (250) مفحوصًا



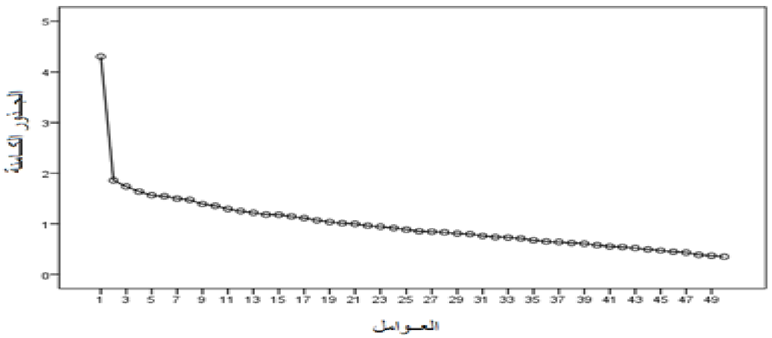
الاختيار من متعدد (25 فقرة)



الاختيار من متعدد (50 فقرة)

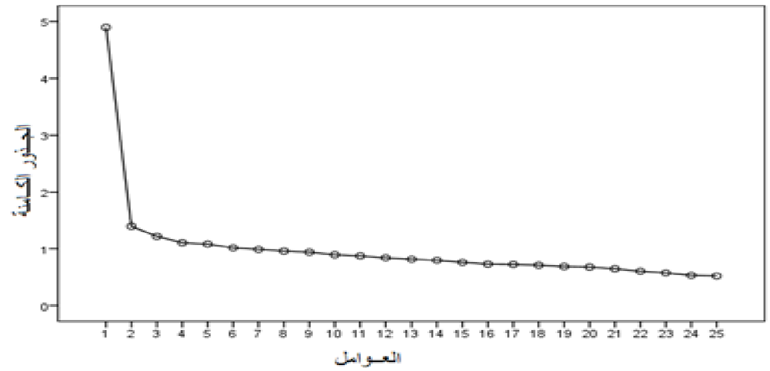


الصواب والخطأ (25 فقرة)

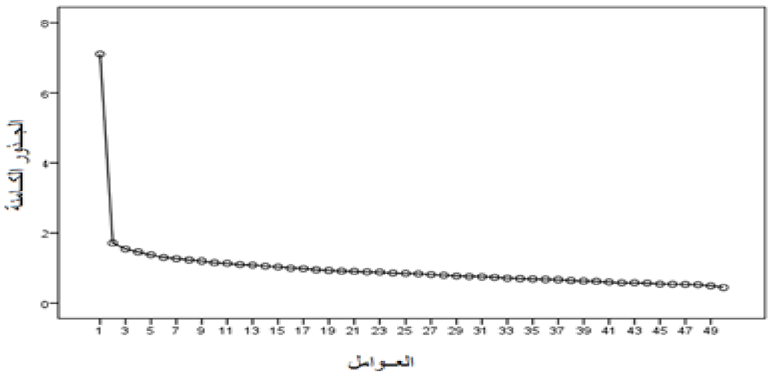


الصواب والخطأ (50 فقرة)

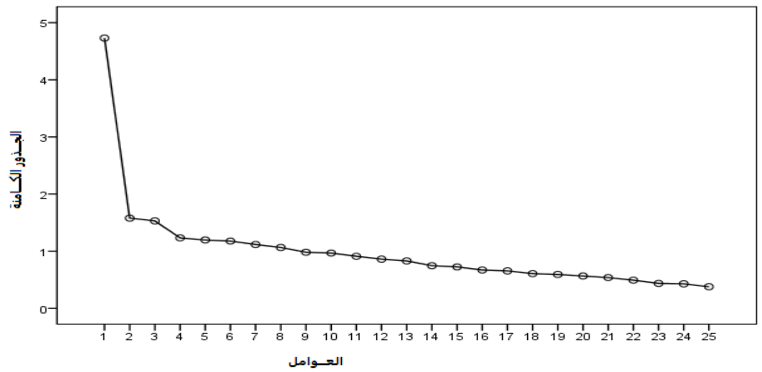
شكل (4): الرسوم البيانية للعلاقة بين قيم الجذور الكامنة والعوامل للإختبارات الأربعة لعينة من المفحوصين حجمها (500) مفحوصًا



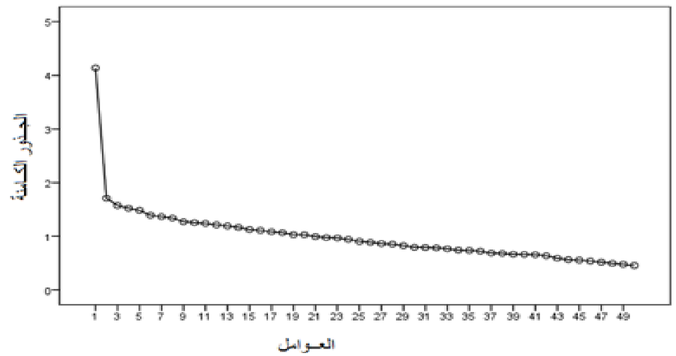
الاختيار من متعدد (25 فقرة)



الاختيار من متعدد (50 فقرة)



الصواب والخطأ (25 فقرة)



الصواب والخطأ (50 فقرة)

الشكل (5): الرسوم البيانية للعلاقة بين قيم الجذور الكامنة والعوامل للإختبارات الأربعة لعينة من المفحوصين حجمها (1015) مفحوصًا

يُلاحظ من خلال الرسوم البيانية الموضحة أعلاه والتي توضح العلاقة البيانية بين العوامل والجذور الكامنة للاختبارات الأربعة وحجوم عيّنات مختلفة مقدار الميل الكبير بين العامل الأول والعامل الثاني مقارنة بالميل بين كل من العامل الثاني والعامل الثالث. ويُلاحظ أيضاً الانحدار الشديد للمنحنى بعد العامل الأول، وإن كان ذلك بنسب متفاوتة تعود إلى نوع الفقرات واختلاف عددها في ضوء اختلاف حجوم العينات، لكن ـ بشكل عام ـ يُعد هذا مؤشرًا على وجود عامل سائد ومسيطر على بقية العوامل في هذه الاختبارات، مما يسمح باستخدام نماذج النظرية الحديثة في القياس.

تمّ حساب معاملات الثبات لكل اختبار، وذلك باستخدام برنامج الرزمة الإحصائية للعلوم الاجتماعية (SPSS) باستخدام معادلة كودر – ريتشاردسون (KR–20)، كما هو موضّح في الجدول (11).

جدول (11)

معامل كودر – ريتشاردسون (KR–20) للاختبارات الأربعة موزّعة حسب حجم العينة

نوع الفقرات		الاختيار من متعدد		الصواب والخطأ	
طول الاختبار		25 فقرة	50 فقرة	25 فقرة	50 فقرة
حجم العينة	250	0.829	0.781	0.746	0.881
	500	0.867	0.877	0.878	0.756
	1015	0.820	0.870	0.801	0.748

يتّضح من الجدول (11) أن قيم معاملات ثبات الاتساق الداخلي للأنواع الأربعة من الاختبار تراوحت بين (0.746) لاختبار الصواب والخطأ القصير في عيّنة حجمها (250) مفحوصًا، و (0.881) لاختبار الصواب والخطأ الطويل في عيّنة حجمها (250) مفحوصًا. ويتّضح أيضاً أن جميع معلاملات الثبات المحسوبة للاختبارات الأربعة، وفي الحجوم الثلاث المختلفة من العيّنات تزيد عن (0.70) وهو الحد الأدنى المقبول لمعامل الثبات في هذا النوع من الاختبارات. وهذا يعني أن معاملات الثبات مقبولة لأغراض الدراسة.

الفصل الرابع

نتائج الدراسة

هدفت هذه الدراسة إلى الكشف عن الأثر الرئيس لكلٍ من متغيرات حجم عينة المفحوصين (250، 500، 1015) مفحوصًا، وعدد فقرات الاختبار (25، 50) فقرة، ونوع الفقرات (أو شكلها) (اختيار من متعدد، صواب وخطأ)، وأثر التفاعلات الثنائية بين هذه المتغيرات، وأثر التفاعل الثلاثي بينها في دقة تقدير معالم الفقرات (الصعوبة، التمييز)، وقدرة الأفراد، ودالة معلومات اختبار تحصيلي في الرياضيات في محتوى الفصل الدراسي الأول من الكتاب المدرسي للصف الثامن الأساسي للعام الدراسي 2016/2015. وقبل الإجابة عن أسئلة الدراسة، تمّ استخراج المتوسطات الحسابية، والانحرافات المعيارية، ودرجة الالتواء، ودرجة التفلطح للدرجات الكليّة الخام على كل اختبار من الاختبارات كما يأتي

جدول (12)

المتوسطات الحسابية والانحرافات المعيارية ودرجة الالتواء ودرجة التفلطح للدرجات الكلية الخام على الاختبارات الأربعة موزّعة حسب متغيّر حجم العيّنة

حجم العينة	نوع الفقرات	عدد فقـــــرات الاختبار							
		50 فقـــرة				25 فقـــرة			
		المتوسط الحسابي	الانحراف المعياري	الالتواء	التفرطح	المتوسط الحسابي	الانحراف المعياري	الالتواء	التفلطح
250	اختيار من متعدد	26.96	6.59	−1.060	1.190	12.86	5.37	0.320	− 0.730
	صواب وخطأ	26.82	9.00	0.009	− 0.482	13.98	4.33	− 0.450	− 0.410
500	اختيار من متعدد	25.86	9.10	0.262	− 0.731	12.64	5.93	0.080	− 0.460
	صواب وخطأ	31.55	6.50	0.212	− 0.247	13.85	6.09	− 1.160	0.780
1015	اختيار من متعدد	26.03	8.80	0.287	− 0.692	12.83	5.26	0.330	− 0.630
	صواب وخطأ	31.32	6.40	0.250	− 0.300	14.71	4.93	− 1.350	2.500

يُلاحظ من الجدول (12) تفاوت قيم المتوسطات الحسابية لدى الأداء على الاختبارات الأربعة، حيث كان متوسط الدرجات الخام التي حصل عليها الطلبة في الاختبار الطويل (50 فقرة) مرتفعاً بشكل عام في اختبارات الصواب والخطأ مقارنة باختبار الاختيار من متعدد وفي جميع مستويات متغيّر حجم العينة؛ وتقاربت قيم المتوسطات الحسابية في اختبار الاختيار من متعدد في جميع حجوم العينة بدرجة كبيرة. أما في الاختبار القصير (25 فقرة)، فقد تبيّن أن هناك ارتفاعاً ملحوظاً في متوسط الدرجات على اختبار الصواب والخطأ أيضاً وفي جميع مستويات حجم العينة، وهناك تقارب في المتوسطات لاختبارات الاختيار من متعدد حيث تزيد بقليل عن (12). أما بالنسبة لقيم الانحراف المعياري، فقد تبيّن أنها بشكل عام أقل في اختبارات الصواب والخطأ الطويلة (50 فقرة) مقارنة باختبارات الاختيار من متعدد الطويلة باستثناء العينة التي كان حجمها (250) مفحوصًا حيث كان الانحراف المعياري لدرجات اختبار الصواب والخطأ أكبر منه في اختبار الاختيار من متعدد. من ناحية ثانية، تبيّن أن قيم الانحراف المعياري للدرجات في اختبارات الصواب والخطأ القصيرة (25 فقرة) أقل منها في اختبارات الاختيار من متعدّد القصيرة باستثناء العينة التي كان حجمها (500) مفحوصًا. ويُلاحظ أيضاً أن قيم الانحراف المعياري للدرجات على الاختبارات القصيرة (25 فقرة) وبغض النظر عن حجم العينة كانت أقل من الانحرافات المعيارية للاختبارات الطويلة (50 فقرة). وعموماً هناك زيادة ملحوظة في قيم الانحراف المعياري عند زيادة عدد الفقرات للاختبار وبغض النظر عن نوع الفقرات وحجم العينة. ولدى النظر في قيم الالتواء يتبيّن أن توزيع الدرجات على الاختبارات المختلفة يقترب من السواء إلى حدٍ ما مع أن هناك أن قيم التواء سالبة وأخرى موجبة في بعض الاختبارات، إلا أن ابتعاد القيم عن القيمة المُثلى (صفر) سواء بالإيجاب أو السلب لم يكن كبيرًا. وكذلك الحال بالنسبة لقيم التفلطح، حيث تبيّن، أن قيم التفرطح للدرجات على الاختبارات المختلفة تقترب من القيمة المُثلى (صفر) باستثناء اختبار الصواب والخطأ القصير (25) فقرة حيث تبيّن أن قيمة التفرطح بلغت (2.50) وهي قيمة مرتفعة إلى حدٍ ما.

وللتأكد فيما إذا كانت الانحرافات الملحوظة لقيم متوسطات درجات الاختبار الخام وانحرافها المعياري تنحرف عن المتوسط الحسابي والانحراف المعياري للتوزيع السويَ، تم استخدام الإحصائي كولمجروف – سمرنوف (Kolmogorov – Smirnova) من خلال الاستعانة بالرزمة الإحصائية للعلوم الاجتماعية؛ لفحص ذلك عند مستوى الدلالة (α=0.05)، (Arnold and Emerson) المشار اليه في (النصراوين، 2015) والجدول (13) الآتي يبين قيم هذا الإحصائي والدلالة الإحصائية.

جدول (13)

نتائج اختبار كولمجروف – سمرنوف لفحص مدى تطابق المتوسط الحسابي والانحراف المعياري للدرجات الخام مع المتوسط الحسابي للتوزيع السوي وانحرافه المعياري تبعاً لحجم العينة وطول الاختبار وشكل فقراته

حجم العينة	نوع الفقرات	عدد فقـــــــرات الاختبار			
		50 فقـــــرة		25 فقـــــرة	
		قيمة الإحصائي (K–S)	مستوى الدلالة	قيمة الإحصائي (K–S)	مستوى الدلالة
250	اختيار من متعدد	0.724	0.416	0.825	0.505
	صواب وخطأ	0.031	0.212	0.826	0.311
500	اختيار من متعدد	0.202	0.334	0.432	0.321
	صواب وخطأ	0.031	0.212	0.452	0.230
1015	اختيار من متعدد	0.234	0.234	0.034	0.201
	صواب وخطأ	0.031	0.212	0.036	0.173

يُلاحظ من الجدول أعلاه أن مستوى الدلالة الإحصائية لجميع قيم اختبار كولمجروف – سمرنوف لبيانات الاختبارات كانت تزيد عن (0.05)، الأمر الذي يشير الى تطابق المتوسط الحسابي والانحراف المعياري لها مع المتوسط الحسابي للتوزيع السويَ والانحراف المعياري له لجميع الاختبارات.

من ناحية ثانية، وقبل إجراء عملية التقدير لمعالم الفقرات والقدرة ودالة معلومات الاختبار كان لا بد من التحقق من حُسن مطابقة فقرات كل اختبار لنموذج التقدير ثنائي المعلمة، ويقصد بمطابقة الفقرات للنموذج هو قدرة أو إمكانية النموذج من تفسير كيفية استجابات الأفراد لإحدى الفقرات أو التنبؤ بهذه

الاستجابات (علام، 2005)، وذلك بإيجاد قيم الإحصائي مربع كاي (χ^2) والاعتماد على مستويات الدلالة، بحيث تكون الفقرة غير مطابقة للنموذج إذا بلغ مستوى الدلالة (0.05) فأقل، وتم اعتماد هذا المؤشر لقوَته وشيوعه في الأدب النظري سواء في إجراء مطابقة الفقرات أو مطابقة الأفراد؛ وتوضح الملاحق (17، 18، 19) قيم الإحصائي (χ^2) ومستوى الدلالة لكل فقرة من فقرات الاختبارات المعتمدة في هذه الدراسة، ولكل مستوى من مستويات حجم العينة الثلاثة (250، 500، 1015) على الترتيب. ويوضّح الجدول (14) أعداد الفقرات التي لم تطابق النموذج الثنائي في الاختبارات الطويلة والقصيرة من النوعين؛ الاختيار من متعدّد، والصواب والخطأ، وفي المستويات الثلاثة من حجم العيّنة.

جدول (14)

الفقرات التي لم تطابق النموذج ثنائي المعلمة المعتمد في التقدير وعددها للاختبارات الأربعة في ضوء حجم العينة

حجم العينة		عدد الفقــرات			
		50 فقــرة		25 فقــرة	
	نوع الفقرات	الفقرات غير المطابقة للنموذج	عددها	الفقرات غير المطابقة للنموذج	عددها
250	اختيار من متعدد	-	0	11	1
	صواب وخطأ	8،34	2	5،7،21	3
500	اختيار من متعدد	-	0	3،11	2
	صواب وخطأ	-	0	3،4،9،25	4
1015	اختيار من متعدد	17،25،28	3	3،4،8،11،13،16	6
	صواب وخطأ	8،9،16،29،31،48	6	3،4،9،11،15،16،19،21،22	9

يتّضح من الجدول (14) أن جميع فقرات اختبار الاختيار من متعدد عندما كان حجم العينة (250) مفحوصًا، وكذلك جميع فقرات كل من اختبار الاختيار من متعدّد واختبار الصواب والخطأ عندما كان حجم العيّنة (500) مفحوصًا في الاختبار الطويل (50 فقرة)، كانت مطابقة للنموذج اللوجستي

ثنائي المعلمة والمعتمد في هذه الدراسة في إجراء عملية التقدير. ويُلاحظ أن أكبر عدد للفقرات غيرالمطابقة للنموذج (تسع فقرات) كان في اختبار الصواب والخطأ القصير (25 فقرة) عندما كان حجم العينة (1015) مفحوصًا. وبما أن الدراسة الحاليّة لم تستهدف تقنين مقياس أو التركيز على الجانب الرياضي الإحصائي والمقارنة بين القيم المقدرة للمعالم، لذا تم اعتماد النتائج دون حذف للفقرات غير المطابقة للنموذج. وفيما يأتي نتائج الإجابة عن أسئلة الدراسة حسب تسلسلها كما وردت في أسئلة الدراسة:

السؤال الأول: "ما الأثر الرئيس لكل متغيّر من متغيّرات حجم عيّنة المفحوصين (250، 500، 1015) مفحوصًا، وطول الاختبار (طويل، قصير)، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ)، وأثر التفاعلات الثنائية الممكنة بين هذه المتغيّرات، والتفاعل الثلاثي بينها في دقة تقدير معالم الفقرات" ؟

للإجابة عن هذا السؤال تمّ الإستعانة بالبرنامج الحاسوبي (Bilog MG) لتقدير معالم صعوبة وتمييز فقرات الاختبارات الأربعة (طويل اختيار من متعدّد، قصير اختيار من متعدّد، طويل صواب وخطأ، قصير صواب وخطأ) وذلك لكل مستوى من مستويات متغّير حجم العيّنة (250، 500، 1015) مفحوصًا على الترتيب كما هو موضّح في الجداول (15، 16، 17).

جدول (15)

قيم تقديرات معالم الفقرات (الصعوبة والتمييز) لكل اختبار من الاختبارات الأربعة لعينة مكوَنة من (250) مفحوصًا

العدد	50 فقرة				25 فقرة				الفقرة	... تابع/ 50 فقرة			
الشكل	اختيار من متعدد		صواب وخطأ		اختيار من متعدد		صواب وخطأ			اختيار من متعدد		صواب وخطأ	
الفقرة	صعوبة	تمييز	صعوبة	تمييز	صعوبة	تمييز	صعوبة	تمييز		صعوبة	تمييز	صعوبة	تمييز
1	2.08−	0.59	1.35−	0.59	0.42−	0.50	0.64−	0.48	**26**	1.46−	0.27	0.73−	0.49
2	0.36	0.28	0.94−	0.29	1.02−	0.60	1.88−	0.48	**27**	0.10	0.37	0.18	0.41
3	1.92−	0.75	1.75−	0.51	0.74−	0.79	1.27−	0.94	**28**	1.54−	0.44	1.05−	0.73
4	2.51−	0.25	2.13−	0.32	0.59−	1.13	1.43−	0.97	**29**	0.63	0.26	0.11	0.50
5	1.83	0.25	0.12	0.34	0.61	0.63	0.14−	0.47	**30**	1.06−	0.44	0.55−	0.66
6	1.23−	0.53	1.74−	0.48	1.77−	0.32	1.72−	0.32	**31**	2.18	0.30	1.34	0.28
7	2.28−	0.51	1.47−	0.58	0.37	0.59	0.55	0.36	**32**	0.42	0.33	0.23	0.36
8	2.16	0.28	0.33	0.45	0.04	0.71	0.40	0.47	**33**	0.26−	0.26	0.19−	0.45
9	3.04−	0.65	3.38−	0.45	0.04−	0.83	0.12−	0.78	**34**	5.28	0.19	2.67	0.21
10	1.52−	0.83	1.12−	0.60	0.30−	0.69	0.47	0.38	**35**	0.09−	0.67	0.42−	0.61
11	0.35	0.50	0.08	0.36	0.48	0.27	1.26	0.33	**36**	0.26	0.33	0.35−	0.38
12	1.49	0.54	0.58	0.28	0.34−	0.90	0.72−	0.41	**37**	0.82	0.23	0.01−	0.38
13	0.51−	0.65	0.79−	0.53	1.09−	0.67	2.04−	0.49	**38**	0.45	0.25	0.09−	0.36
14	0.15	0.67	0.07	0.40	0.40	0.87	0.21	0.41	**39**	0.41−	0.39	0.43−	0.42
15	1.08−	0.30	1.01−	0.50	0.16−	0.87	0.98−	0.98	**40**	1.03	0.34	0.09−	0.36
16	1.56−	0.59	0.93−	0.68	0.49	0.52	0.09−	0.62	**41**	0.08	0.28	0.51	1.46
17	1.24	0.38	0.21−	0.33	0.15−	0.67	1.14−	1.25	**42**	0.22	0.52	0.42	1.68
18	0.66−	0.42	0.92−	0.46	0.36	0.55	0.85	0.53	**43**	0.12−	0.37	0.62	0.97
19	1.66	0.37	0.34	0.49	0.49−	0.57	0.89	0.61	**44**	1.62−	0.75	0.17	1.77
20	1.08	0.49	0.23	0.40	0.87	0.51	1.55	0.41	**45**	1.41	0.34	0.82	1.08
21	0.88	0.51	0.04−	0.54	0.16	0.66	0.57	0.51	**46**	0.61−	0.31	0.51	1.36
22	1.64−	0.32	1.46−	0.37	0.16−	0.69	0.20−	0.57	**47**	0.58−	0.51	0.35	1.32
23	1.54−	0.53	1.12−	0.55	0.52−	0.98	1.41−	0.59	**48**	3.25	0.30	1.62	0.63
24	0.75−	0.56	0.35−	0.74	1.38	0.35	1.81	0.29	**49**	1.28−	0.40	0.41	1.21
25	1.32−	0.54	0.81−	0.65	0.61	0.68	2.05	0.29	**50**	1.62	0.29	0.75	0.70

يُلاحظ من الجدول أعلاه أن قيم الصعوبة لجميع الاختبارات كانت ضمن الفترة الواقعية (−3،3) لوجت، باستثناء الفقرات (9، 34، 48) في اختبار الاختيار من متعدد الطويل، والفقرة (9) من اختبار الصواب والخطأ الطويل، ويبدو التدني في قدرات المفحوصين بدليل كثرة التقديرات السالبة لمعلمة

الصعوبة، وبالنسبة لمعلمة التمييز فإن جميع تقديراتها موجبة ومقبولة لجميع الاختبارات، مما يعني القدرة الجيدة لهذه الفقرات على التمييز بين الأفراد.

جدول (16)

قيم تقديرات معالم الفقرات (الصعوبة والتمييز) لكل اختبار من الاختبارات الأربعة لعينة مكوَنة من (500) مفحوصًا

العدد	50 فقرة				25 فقرة				الفقرة	... تابع/ 50 فقرة			
الشكل	اختيار من متعدد		صواب وخطأ		اختيار من متعدد		صواب وخطأ			اختيار من متعدد		صواب وخطأ	
الفقرة	صعوبة	تمييز	صعوبة	تمييز	صعوبة	تمييز	صعوبة	تمييز		صعوبة	تمييز	صعوبة	تمييز
1	−0.76	0.83	−1.16	0.37	−0.21	0.61	−0.77	1.30	**26**	−0.36	0.41	−1.77	0.25
2	−0.30	0.45	−1.36	0.24	−0.76	0.86	−0.52	1.12	**27**	−0.07	0.40	−1.05	0.29
3	−1.09	0.58	−1.34	0.41	−0.62	1.13	−0.14	1.73	**28**	−0.71	0.63	−1.74	0.45
4	−1.83	0.28	−2.04	0.39	−0.51	1.16	−0.12	1.58	**29**	−0.16	0.39	0.68	0.42
5	0.55	0.44	−0.58	0.49	0.68	0.53	0.78	0.54	**30**	−0.16	0.85	−1.38	0.61
6	−1.50	0.46	−1.81	0.59	−1.03	0.72	−0.55	0.94	**31**	0.98	0.41	−0.69	0.53
7	−1.69	0.33	−1.14	0.45	0.37	0.74	−0.20	0.74	**32**	0.18	0.61	−0.05	0.27
8	1.02	0.36	0.86	0.23	0.19	0.80	0.47	0.67	**33**	−0.07	0.72	−1.33	0.31
9	−3.07	0.48	−2.01	1.02	−0.02	1.04	−0.47	1.20	**34**	4.99	0.17	0.80	0.20
10	−1.22	0.46	−0.86	0.59	−0.17	1.02	−0.13	0.76	**35**	0.08	0.75	−1.07	0.53
11	0.58	0.50	−0.59	0.47	0.42	0.35	0.48	0.49	**36**	−0.20	0.69	−0.75	0.43
12	0.82	0.42	0.09	0.37	−0.26	0.80	−0.37	1.16	**37**	0.07	0.37	0.05	0.25
13	−0.48	0.56	−1.64	0.58	−0.84	0.99	−0.64	1.32	**38**	0.13	0.74	−1.38	0.26
14	0.84	0.43	−0.08	0.29	0.38	0.78	0.42	0.52	**39**	−0.23	0.89	−0.15	0.37
15	−0.17	0.51	−1.55	0.51	−0.02	1.31	−0.08	0.88	**40**	−0.13	0.60	−1.23	0.22
16	−0.69	0.68	−1.17	0.46	0.37	0.64	−0.11	0.61	**41**	0.20	0.57	−1.21	0.47
17	1.10	0.30	−1.17	0.24	−0.12	1.04	−0.47	1.08	**42**	0.09	0.74	−0.94	0.51
18	−0.70	0.46	−0.81	0.43	0.40	0.60	−0.05	0.67	**43**	0.09	0.41	−0.37	0.26
19	0.65	0.50	−0.80	0.21	−0.32	0.61	−0.31	0.78	**44**	−0.63	0.62	−1.11	0.53
20	0.83	0.55	−0.47	0.25	0.64	0.62	0.53	0.59	**45**	0.62	0.48	0.08	0.28
21	0.48	0.65	0.17	0.41	−0.04	0.81	−0.15	0.91	**46**	−0.33	0.63	−1.00	0.59
22	−0.72	0.51	−1.92	0.45	0.01	0.80	−0.11	0.94	**47**	−0.35	0.60	−0.72	0.45
23	0.56	0.61	−1.10	0.66	−0.23	1.12	0.37	0.74	**48**	1.61	0.30	−0.41	0.23
24	−0.04	0.44	1.11	0.60	1.30	0.35	0.10	0.70	**49**	0.06	0.48	−2.51	0.25
25	−0.40	0.75	−0.91	0.56	0.78	0.79	0.71	0.87	**50**	0.32	0.35	−1.22	0.46

يُلاحظ من الجدول أعلاه أن قيم الصعوبة لجميع الاختبارات كانت ضمن الفترة الواقعية (−3،3) لوجت، باستثناءالفقرتين (9، 34) في اختبار الاختيار من متعدد الطويل، ويبدو التدني في قدرات

المفحوصين بدليل كثرة التقديرات السالبة لمعلمة الصعوبة. وبالنسبة لمعلمة التمييز فإن جميع تقديراتها موجبة ومقبولة لجميع الاختبارات، مما يعني القدرة الجيدة لهذه الفقرات على التمييز بين الأفراد.

جدول (17)

قيم تقديرات معالم الفقرات (الصعوبة والتمييز) لكل اختبار من الاختبارات الأربعة لعينة مكوَنة من (1015) مفحوصًا

العدد	50 فقرة				25 فقرة				الفقرة	... تابع 50/ فقرة			
الشكل	اختيار من متعدد		صواب وخطأ		اختيار من متعدد		صواب وخطأ			اختيار من متعدد		صواب وخطأ	
الفقرة	صعوبة	تمييز	صعوبة	تمييز	صعوبة	تمييز	صعوبة	تمييز		صعوبة	تمييز	صعوبة	تمييز
1	−0.85	0.68	−1.20	0.34	−0.34	0.45	−1.35	0.91	**26**	−0.51	0.40	−2.18	0.19
2	−0.36	0.37	−1.31	0.23	−1.10	0.65	−1.11	0.69	**27**	0.07	0.39	−1.05	0.24
3	−1.10	0.56	−1.38	0.43	−0.74	0.97	−0.47	1.24	**28**	−0.70	0.69	−1.80	0.43
4	−2.18	0.23	−2.08	0.37	−0.67	1.02	−0.35	1.25	**29**	−0.17	0.32	0.85	0.38
5	0.64	0.36	−0.59	0.45	0.84	0.41	0.96	0.35	**30**	−0.24	0.75	−1.37	0.60
6	−1.36	0.46	−1.95	0.50	−1.57	0.40	−1.07	0.62	**31**	1.05	0.36	−0.47	0.60
7	−1.76	0.37	−1.14	0.42	0.42	0.57	−0.59	0.49	**32**	0.21	0.60	−0.27	0.28
8	1.09	0.36	1.17	0.19	0.25	0.64	0.50	0.44	**33**	−0.04	0.61	−1.07	0.34
9	−3.26	0.42	−2.13	0.89	−0.07	0.86	−0.83	0.96	**34**	5.73	0.13	1.00	0.18
10	−1.14	0.46	−0.82	0.54	−0.31	0.72	−0.48	0.48	**35**	0.08	0.68	−1.06	0.50
11	0.62	0.40	−0.65	0.41	0.38	0.28	0.61	0.23	**36**	−0.33	0.63	−0.86	0.43
12	0.92	0.39	0.35	0.28	−0.36	0.66	−0.73	0.82	**37**	0.06	0.44	0.18	0.22
13	−0.44	0.54	−1.77	0.51	−0.96	0.86	−1.11	0.96	**38**	0.09	0.62	−01.4	0.28
14	0.82	0.43	−0.03	0.29	0.42	0.63	0.53	0.29	**39**	−0.21	0.74	−0.06	0.34
15	−0.18	0.51	−1.64	0.44	−0.07	0.90	−0.33	0.68	**40**	−0.10	0.64	−1.28	0.18
16	−0.67	0.68	−1.10	0.48	0.52	0.42	−0.39	0.38	**41**	0.22	0.55	−1.26	0.43
17	0.94	0.33	−1.77	0.20	−0.20	0.82	−0.89	0.73	**42**	0.04	0.73	−1.07	0.44
18	−0.91	0.45	−0.93	0.42	0.58	0.46	−0.36	0.47	**43**	0.03	0.50	−0.13	0.26
19	0.65	0.47	−0.73	0.18	−0.44	0.43	−0.74	0.53	**44**	−0.56	0.72	−1.80	0.48
20	0.83	0.42	−0.30	0.23	0.89	0.43	0.65	0.37	**45**	0.63	0.49	−0.08	0.21
21	0.44	0.60	0.24	0.38	−0.06	0.68	−0.42	0.62	**46**	−0.24	0.60	−0.99	0.55
22	−0.74	0.51	−1.86	0.44	−0.03	0.63	−0.36	0.69	**47**	−0.30	0.57	−0.96	0.38
23	−0.70	0.54	−1.10	0.58	−0.27	0.86	0.28	0.54	**48**	0.67	0.29	−0.32	0.19
24	−0.12	0.45	0.06	0.50	1.68	0.27	−0.05	0.47	**49**	0.08	0.39	−2.88	0.23
25	0.56	0.73	−0.83	0.59	0.87	0.52	0.79	0.63	**50**	0.44	0.36	−1.43	0.42

يُلاحظ من الجدول أعلاه أن قيم الصعوبة لجميع الاختبارات كانت ضمن الفترة الواقعية (−3،3) لوجت، باستثناء الفقرتين (9، 34) في اختبار الاختيار من متعدد الطويل، ويبدو التحسن في قدرات

المفحوصين بدليل انخفاض عدد التقديرات السالبة لمعلمة الصعوبة. وبالنسبة لمعلمة التمييز فإن جميع تقديراتها موجبة ومقبولة لجميع الاختبارات، مما يعني القدرة الجيدة لهذه الفقرات على التمييز بين الأفراد.

بعد ذلك تمّ استخراج المتوسطات الحسابية والانحرافات المعيارية لتقديرات معالم الصعوبة والتمييز لفقرات كل اختبار من الاختبارات الأربعة موزّعة حسب مستويات متغيّر حجم عيّنة المفحوصين كما يوضّحها الجدول الآتي:

جدول (18)

المتوسطات الحسابية (م) والانحرافات المعيارية (ع) لتقديرات معالم الفقرات لكل اختبار من الاختبارات الأربعة موزّعة حسب مستويات متغيّر حجم عيّنة المفحوصين

حجم العيّنة	شكل الاختبار	طول الاختبار							
		طويل (50 فقرة)				قصير (25 فقرة)			
		الصعوبة		التمييز		الصعوبة		التمييز	
		م	ع	م	ع	م	ع	م	ع
250	**اختيار من متعدّد**	−0.074	1.578	0.428	0.158	−0.081	0.690	0.661	0.201
	صواب وخطأ	−0.250	1.010	0.613	0.365	−0.198	1.165	0.557	0.247
500	**اختيار من متعدّد**	−0.046	1.101	0.526	0.158	0.019	0.555	0.808	0.246
	صواب وخطأ	−0.875	0.751	0.414	0.157	−0.054	0.432	0.913	0.328
1015	**اختيار من متعدّد**	−0.047	1.203	0.498	0.144	−0.014	0.721	0.622	0.214
	صواب وخطأ	−0.891	0.848	0.381	0.147	−0.292	0.659	0.632	0.269

يُلاحظ من الجدول (18) أن متوسط تقديرات معاملات صعوبة فقرات اختبارات الاختيار من متعدّد أكبر منها بالنسبة لفقرات اختبارات الصواب والخطأ سواء للاختبارات الطويلة (50 فقرة) أو

الاختبارات القصيرة (25 فقرة)، وبصرف النظر عن حجم العيّنة. ويُلاحظ أيضاً أن متوسط تقديرات معاملات تمييز فقرات اختبارات الاختيار من متعدّد أكبر من معاملات تمييز فقرات اختبار الصواب والخطأ عندما يكون حجم العيّنة (500) أو (1015) في الاختبارات الطويلة، بينما يزيد متوسط معاملات تمييز فقرات اختبار الصواب والخطأ عنه بالنسبة لفقرات اختبار الاختيار من متعدّد في هذه الاختبارات عندما يكون حجم العيّنة (250) مفحوصًا. أما في الاختبارات القصيرة، فقد تبيّن أن متوسط معاملات تمييز فقرات الاختبارات القصيرة أكبر منها بالنسبة للاختبارات نفسها من نوع الاختيار من متعدّد عندما يكون حجم العيّنة (500) مفحوصًا أو (1015) مفحوصًا، بينما يزيد متوسط معاملات تمييز اختبارات الاختيار من متعدّد القصيرة عنه في اختبارات الصواب والخطأ القصيرة. وبالنسبة لقيم لانحراف المعياري لمعاملات الصعوبة، فيُلاحظ أنها أكبر في اختبارات الاختيار من متعدّد مقارنة باختبارات الصواب والخطأ الطويلة بصرف النظر عن حجم العيّنة، وكذلك الحال في الاختبارات القصيرة باستثناء الحالة التي يكون فيها حجم العيّنة (250) مفحوصًا. وبالمثل، تبيّن أن قيم الانحراف المعياري لمعاملات تمييز اختبارات الاختيار من متعدّد الطويلة تزيد عنها زيادة ملحوظة مقارنة باختبارات الصواب والخطأ عندما يكون حجم العيّنة (250) مفحوصًا، وتتقارب هذه القيم من بعضها بعضًا عندما يزداد حجم العيّنة. وفي الاختبارات القصيرة، تبيّن أن قيم الانحراف المعياري لمعاملات التمييز يزداد بزيادة حجم العيّنة.

وللكشف عن دقة تقدير معالم الفقرات (الصعوبة، والتمييز) تمّ إيجاد متوسطات الخطأ المعياري للتقدير Standard Error of Estimate (SEE) في ظل متغيّرات حجم العيّنة، وطول الاختبار، ونوع الاختبار، حيث يُعدّ هذا المؤشر الإحصائي من مؤشرات قياس دقة تقدير معالم الفقرات والأفراد ضمن نظرية الاستجابة للفقرة؛ فكلما قلّت قيمة هذا المؤشر زادت دقة تقدير المعالم. ويوضح جدول (19) قيم

متوسطات الخطأ المعياري لتقدير معالم الفقرات (الصعوبة، والتمييز) باختلاف حجم العينة وعدد فقرات الاختبار ونوع هذه الفقرات.

جدول (19)

المتوسطات الحسابية لقيم الخطأ المعياري لتقدير معالم الفقرات

حجم العيّنة	نوع الاختبار	طول الاختبار			
		50 فقرة		25 فقرة	
		الصعوبة	التمييز	الصعوبة	التمييز
250	اختيار من متعدّد	0.079	0.347	0.097	0.170
	صواب وخطأ	0.087	0.226	0.088	0.249
500	اختيار من متعدّد	0.061	0.173	0.080	0.100
	صواب وخطأ	0.064	0.226	0.110	0.104
1015	اختيار من متعدّد	0.044	0.140	0.051	0.096
	صواب وخطأ	0.046	0.185	0.064	0.111

يُلاحظ من الجدول أعلاه أن قيم متوسطات الخطأ المعياري لتقدير معالم الفقرات تختلف باختلاف متغيّرات حجم العيّنة، وطول الاختبار، وشكل الفقرات. ويبيّن الجدول أن أقل قيمة بين قيم متوسطات الخطأ المعياري لتقدير معلمة الصعوبة بلغت (0.044) لاختبار الاختيار من متعدّد الطويل عندما كان حجم العيّنة (1015) مفحوصًا، بينما كانت أكبر قيمة بين تلك القيم (0.110) لاختبار الصواب والخطأ القصير عندما كان حجم العيّنة (500). وبالنسبة لتقدير معلمة التمييز فإن أقل قيمة لمتوسطات الخطأ المعياري في تقديرها بلغت (0.096) للإختبار القصير فقراته من نوع الاختيار من متعدد وعند الحجم الأكبر للعينة، بينما بلغت أكبر قيمة (0.347) للإختبار الطويل الذي فقراته من نوع اختيار من متعدد ولحجم العينة الأصغر.

وللكشف عن دلالة الفروق بين قيم متوسطات الأخطاء المعيارية لتقدير معالم الفقرات باختلاف مستويات متغيّرات طول الاختبار (طويل، قصير)، وحجم عيّنة المفحوصين (250، 500، 1015) مفحوصًا، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ) تمّ استخدام أسلوب تحليل التباين العاملي (2x3x2) ذو القياسات المتكرّرة repeated measures factorial ANOVA أحادي بين الأفراد وثنائي داخل الأفراد (One Between – Two Within Subjects Design)، حيث يمثّل متغيّر طول الاختبار (طويل، قصير) متغيّر بين الأفراد، بينما يمثّل متغيّري حجم العيّنة (250، 500، 1015)، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ) متغيّري داخل الأفراد. ولأغراض تسهيل قراءة النتائج، تمّ تقسيمها إلى جزأين؛ تناول الجزء الأول معلمة الصعوبة، فيما تناول الجزء الثاني معلمة التمييز كما يأتي:

(أ) النتائج المتعلقة بتقديرات معلمة الصعوبة: للكشف عن الآثار الرئيسة للمتغيّرات المستقلة (طول الاختبار، حجم العيّنة، شكل الاختبار)، وأثر التفاعلات الثنائية الممكنة بينها، وأثر التفاعل الثلاثي في تقدير معلمة الصعوبة تمّ حساب المتوسطات الحسابية والانحرافات المعيارية لقيم الخطأ المعياري لتقدير معلمة الصعوبة لكل اختبار من الاختبارات الأربعة (اختيار من متعدّد طويل، اختيار من متعدّد قصير، صواب وخطأ طويل، صواب وخطأ قصير) ولكل عيّنة من العيّنات الثلاث (250، 500، 1015) كما يوضّحها الجدول (20) الآتي.

جدول (20)

المتوسطات الحسابية (م) والانحرافات المعيارية (ع) لقيم الأخطاء المعيارية لتقدير معلمة الصعوبة تبعاً لمتغيرات شكل الاختبار وحجم العينة وطول الاختبار (ن)

شكل الاختبار	حجم العيّنة	طول الاختبار	ن	م	ع
اختيار من متعدّد	250	طويل	50	0.079	0.015
		قصير	25	0.097	0.017
		المجموع	75	0.085	0.018
	500	طويل	50	0.061	0.008
		قصير	25	0.080	0.012
		المجموع	75	0.067	0.013
	1015	طويل	50	0.044	0.005
		قصير	25	0.051	0.008
		المجموع	75	0.046	0.007
صواب وخطأ	250	طويل	50	0.087	0.020
		قصير	25	0.088	0.019
		المجموع	75	0.087	0.019
	500	طويل	50	0.064	0.016
		قصير	25	0.110	0.018
		المجموع	75	0.079	0.027
	1015	طويل	50	0.046	0.010
		قصير	25	0.064	0.013
		المجموع	75	0.052	0.014

عند مقارنة مجاميع المتوسطات الحسابية لقيم الأخطاء المعيارية لاختبارات اختيار من متعدد كان المجموع الأقل عند العينة الأكبر حيث بلغ (0.046)، وبالمثل كان مجموع المتوسطات الأقل لاختبارات الصواب والخطأ عند العينة الأكبر حيث بلغ (0.052). ويتّضح من الجدول (20) وجود فروق ظاهرية بين المتوسطات الحسابية لقيم الأخطاء المعيارية لتقدير معلمة الصعوبة. وللكشف عن هذه الفروق تمّ استخدام أسلوب تحليل التباين العاملي (2x3x2) ذو القياسات المتكرّرة (أحادي بين الأفراد وثنائي داخل الأفراد One Between – Two Within Subjects) كما يوضّحها الجدول (21).

جدول (21)

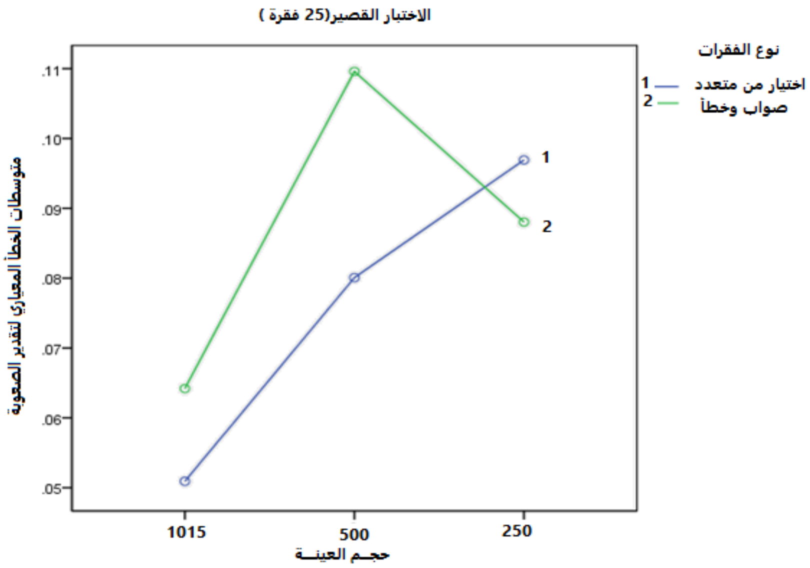
نتائج تحليل التباين العاملي (2x3x2) ذو القياسات المتكرّرة للكشف عن أثر المتغيّرات المستقلة في تقدير معلمة الصعوبة

مصدر التباين	م. م	د.ح	م. م. م	قيمة "ف"	η^2 الجزئي
بين الأفراد					
طول الاختبار (أ)	0.033	1	0.033	52.170**	0.417
داخل الأفراد					
شكل الاختبار (ب)	0.006	1	0.006	37.321**	0.338
حجم العيّنة (ج)	0.096	2	0.048	518.458**	0.877
أ x ب	0.001	1	0.001	6.901*	0.086
أ x ج	0.010	2	0.005	55.721**	0.433
ب x ج	0.005	2	0.002	24.357**	0.250
أ x ب x ج	0.008	2	0.004	40.803**	0.359

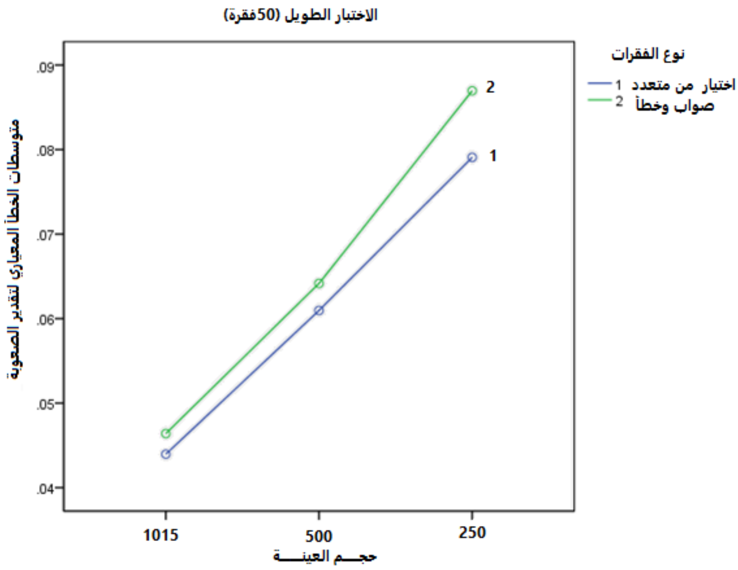
* $\alpha \geq 0.05$ (حيث تدل م.م : على مجموع المربعات، د.ح : درجات الحرية، م.م.م : معدل مجموع المربعات)

** $\alpha \geq 0.01$

يتّضح من الجدول (21) أن لتفاعل طول الاختبار وشكل الاختبار أثر دال إحصائياً على قيم متوسطات الخطأ المعياري في تقدير معلمة الصعوبة وبالتالي على دقة التقدير، وفسر هذا التفاعل مانسبته 8.6% من التباين في قيمة الخطأ المعياري؛ أما تفاعل طول الاختبار وحجم العينة فكان له أثر دال احصائياً وفسر ما نسبته 43.3% من تباين الخطأ المعياري، أما التفاعل بين شكل الاختبار وحجم العينة كان أثره دال احصائياً أيضاً وفسر 2.5% من تباين الخطأ المعياري. وهناك أثر دال للتفاعل الثلاثي بين المتغيرات الثلاثة وفسر ما نسبته 35.9% من تباين الخطأ المعياري في تقدير الصعوبة كما توضّحها الأشكال (6، 7).



شكل (6): أثر التفاعل بين متغيّري شكل الاختبار(نوع الفقرات) وحجم العيّنة للاختبار القصير (25 فقرة) في قيمة الخطأ المعياري لتقدير معلمة الصعوبة



شكل (7): أثر التفاعل بين متغيّري شكل الاختبار وحجم العيّنة للاختبار الطويل (50 فقرة) في قيمة الخطأ المعياري لتقدير معلمة الصعوبة

يتبيّن من الشكلين (6، 7) أن متوسطات الخطأ المعياري في تقدير معلمة الصعوبة في اختبارات الاختيار من متعدّد القصيرة والطويلة على حد سواء أقل منها مقارنة باختبارات الصواب والخطأ وبصرف

النظر عن حجم العيّنة، إلا أن الفرق بين قيم الخطأ المعياري في هذه الاختبارات يظهر بوضوح عندما يزداد حجم العيّنة.

من ناحية ثانية، أظهر التحليل وجود أثر رئيس لكل متغيّر من المتغيّرات المستقلة الثلاثة (طول الاختبار، حجم العيّنة، شكل الاختبار) في المتغيّر التابع (الخطأ المعياري لتقدير معلمة الصعوبة)؛ فبالنسبة لمتغيّر الطول تبيّن أن قيم الخطأ المعياري لتقدير معلمة الصعوبة في الاختبارات الطويلة مهما كان شكلها (اختيار من متعدّد، صواب وخطأ) وبصرف النظر عن حجم العيّنة أقل منها مقارنة بالاختبارات القصيرة، أما بالنسبة لمتغيّر شكل الاختبار، فقد تبيّن أن قيم الخطأ المعياري لتقدير معلمة الصعوبة في اختبارات الاختيار من متعدّد أقل منها إذا ما قورنت باختبارات الصواب والخطأ وبصرف النظر عن طول الاختبار أو حجم العيّنة، وبالنسبة لمتغيّر حجم العيّنة، فقد أظهر التحليل أن الخطأ المعياري لتقدير معلمة الصعوبة يقل كلما ازداد حجم العيّنة.

من جانبٍ آخر، كشف التحليل أن قيم حجم الأثر الجزئي (Partial η^2) للأثر الرئيس لكل متغيّر من المتغيّرات المستقلة الثلاثة (طول الاختبار، حجم العيّنة، شكل الاختبار)، وأثر التفاعلات الممكنة بين هذه المتغيّرات تزيد عن (0.13)، وهي قيم كبيرة بحسب معيار كوهن Cohen لحجم الأثر باستثناء أثر التفاعل بين متغيّري طول الاختبار وشكل الاختبار الذي كان متوسطًا حيث بلغت قيمته (0.086) وهو يزيد عن (0.06) لكنه يقل عن (0.13) أي أن حجم الأثر متوسط. كما أظهر التحليل أن متغيّر حجم العيّنة فسّر أكبر نسبة من التباين في المتغيّر التابع (قيمة الخطأ المعياري لتقدير معلمة الصعوبة) مقارنة بالأثر الرئيس لكل من متغيّري طول الاختبار وشكل الاختبار حيث بلغت قيمة (Partial η^2) لهذا المتغيّر (0.877)، مما يعني أن متغيّر حجم العيّنة فسّر ما مقداره (% 87.7) من التباين في قيمة الخطأ المعياري، وكذلك بالنسبة لأثر كل نوع من أنواع التفاعلات المختلفة.

ولمعرفة مصدر الفروق الإجمالية الذي كشف عنه تحليل التباين للأثر الرئيس لمتغيّر حجم العيّنة فقد تمّ استخدام اختبار "بونفيوروني" Bonferroni للمقارنات البعدية كما في الجدول الآتي.

جدول (22)

نتائج اختبار "بونفيوروني" Bonferroni للمقارنات البعدية للكشف عن مصدر التباين للأثر الرئيس لمتغيّر حجم العيّنة في قيمة الخطأ المعياري لتقدير معلمة الصعوبة

المتوسط الحسابي	حجم العيّنة	500	1015
0.088	250	0.009*	0.037*
0.079	500		0.028*
0.051	1015		

* α = 0.05

يتّضح من الجدول (22) أن الفروق الثلاثة بين متوسطات قيم الخطأ المعياري لتقدير معلمة الصعوبة في الحجوم المختلفة للعيّنة كانت ذات دلالة إحصائية عند مستوى دلالة يقل عن (0.05).

(ب) النتائج المتعلقة بتقديرات معلمة التمييز: للكشف عن الآثار الرئيسة للمتغيّرات المستقلة (طول الاختبار، حجم العيّنة، شكل الاختبار)، وأثر التفاعلات الثنائية الممكنة بينها، وأثر التفاعل الثلاثي في تقدير معلمة التمييز تمّ حساب المتوسطات الحسابية والانحرافات المعيارية لقيم الخطأ المعياري لتقدير معلمة التمييز لكل اختبار من الاختبارات الأربعة (اختيار من متعدّد قصير، اختيار من متعدّد طويل، صواب وخطأ قصير، صواب وخطأ طويل) ولكل عيّنة من العيّنات الثلاث (250، 500، 1015) كما يوضّحها الجدول (23).

جدول (23)

المتوسطات الحسابية (م) والانحرافات المعيارية (ع) لقيم الأخطاء المعيارية لتقدير معلمة التمييز تبعاً لمتغيرات شكل الاختبار وحجم العينة وطول الاختبار (ن)

شكل الاختبار	حجم العيّنة	طول الاختبار	ن	م	ع
اختيار من متعدّد	250	قصير	25	0.170	0.043
		طويل	50	0.347	0.015
		المجموع	75	0.086	0.029
	500	قصير	25	0.100	0.043
		طويل	50	0.173	0.181
		المجموع	75	0.149	0.153
	1015	قصير	25	0.096	0.052
		طويل	50	0.140	0.186
		المجموع	75	0.125	0.156
صواب وخطأ	250	قصير	25	0.249	0.118
		طويل	50	0.226	0.144
		المجموع	75	0.233	0.136
	500	قصير	25	0.104	0.024
		طويل	50	0.226	0.090
		المجموع	75	0.185	0.094
	1015	قصير	25	0.111	0.035
		طويل	50	0.185	0.091
		المجموع	75	0.160	0.084

يتّضح من الجدول أعلاه أن المجموع الأقل للمتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التمييز لاختبارات اختيار من متعدد كان عند حجم العينة الأقل حيث بلغ (0.086)، بينما كان المجموع الأقل لفقرات صواب وخطأ عند العينة الأكبر إذ بلغ (0.160). ووجود الفروق الظاهرية بين المتوسطات الحسابية لقيم الأخطاء المعيارية لتقدير معلمة التمييز. وللكشف عن هذه الفروق تمّ استخدام أسلوب تحليل التباين العاملي (2x3x2) ذو القياسات المتكرّرة (أحادي بين الأفراد وثنائي داخل الأفراد One Between – Two Within Subjects) كما يوضّحها الجدول (24).

جدول (24)

نتائج تحليل التباين العاملي (2x3x2) ذو القياسات المتكرّرة للكشف عن أثر المتغيّرات المستقلة في تقدير معلمة التمييز

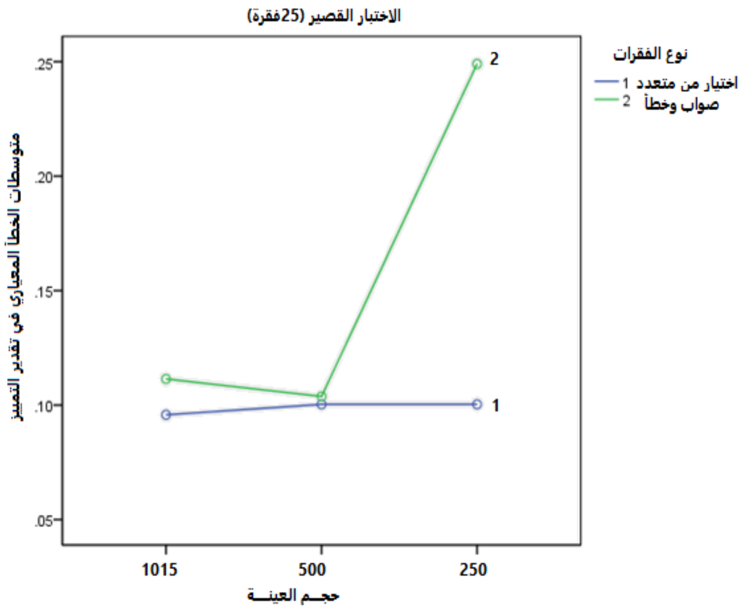
مصدر التباين	م. م	د.ح	م. م. م	قيمة "ف"	η^2 الجزئي
بين الأفراد	8.895	1	8.895	**237.889	0.765
طول الاختبار (أ)	0.200	1	0.200	*5.340	0.068
داخل الأفراد					
شكل الاختبار (ب)	0.471	1	0.471	**46.553	0.389
حجم العيّنة (ج)	0.063	2	0.032	**8.885	0.109
أ x ب	0.016	1	0.016	1.587	0.021
أ x ج	0.249	2	0.125	**35.116	0.325
ب x ج	0.313	2	0.156	**13.570	0157
أ x ب x ج	0.011	2	0.005	0.476	0.006

* $\alpha \leq 0.05$ (حيث تدل م.م : على مجموع المربعات، د.ح : درجات الحرية، م.م.م : معدل مجموع المربعات)

** $\alpha \leq 0.01$

يتّضح من الجدول (24) أعلاه وجود أثر لجميع التفاعلات الثنائية المحتملة بين المتغيّرات المستقلة في المتغيّر التابع باستثناء تفاعل متغيّري طول الاختبار مع شكل الاختبار، والتفاعل الثلاثي بين متغيّرات طول الاختبار وحجم العيّنة وشكل الاختبار. وفيما يتعلق بحجم الأثر للتفاعلات بين المتغيّرات الثلاثة، فقد كشف التحليل عن وجود حجم أثر كبير لتفاعل متغيّر حجم العيّنة مع كلٍ من متغيّري طول الاختبار وشكل الاختبار زادت عن معيار حجم الأثر الكبير بحسب معيار "كوهن" Cohen وهو (0.13)، في حين تبيّن أن لتفاعل متغيّري طول الاختبار وشكل الاختبار أثر صغير في تفسير التباين في قيمة الخطأ المعياري لتقدير معلمة التمييز لم تتجاوز (2%). من ناحية أخرى لم

يتبين وجود أثر لتفاعل المتغيّرات الثلاثة (طول الاختبار، حجم العيّنة، شكل الاختبار) في تفسير التباين في قيمة الخطأ المعياري لتقدير معلمة التمييز. من جانبٍ آخر، كشف التحليل عن أن قيم حجم الأثر الجزئي (Partial η^2) للأثر الرئيس لكل متغيّر من المتغيّرات المستقلة الثلاثة (طول الاختبار، حجم العيّنة، شكل الاختبار) تراوحت بين حجم الأثر الكبير والمتوسط، حيث أظهرت النتائج وجود حجم أثر كبير لمتغيّر شكل الاختبار (اختيار من متعدّد، صواب وخطأ) في تباين قيمة الخطأ المعياري لمعلمة التمييز، وفسّر عامل شكل الاختبار ما نسبته (38.90%) من التباين في المتغيّر التابع (الخطأ المعياري لتقدير معلمة التمييز)، في حين كانت قيم حجم الأثر لكل من متغيّري حجم العيّنة وطول الاختبار متوسطة، حيث فسّر كل منهما ما نسبته (10.90%، 6.80%) على الترتيب من التباين في المتغيّر التابع (الخطأ المعياري لتقدير معلمة التمييز). وتوضّح الأشكال (8، 9) أثر التفاعلات التي كشف عنها التحليل.



شكل (8): أثر التفاعل بين متغيّري شكل الاختبار وحجم العيّنة للاختبار القصير (25 فقرة) في قيمة الخطأ المعياري لتقدير معلمة التمييز



شكل (9): التفاعل بين متغيّري شكل الاختبار وحجم العيّنة للاختبار الطويل (50 فقرة) في قيمة الخطأ المعياري لتقدير معلمة التمييز

يتبيّن من الشكلين (8، 9) أن الخطأ المعياري في تقدير معلمة التمييز في اختبارات الاختيار من متعدّد القصيرة والطويلة على حد سواء أقل منها مقارنة باختبارات الصواب والخطأ وبصرف النظر عن حجم العيّنة، إلا أن الفرق بين قيم الخطأ المعياري في هذه الاختبارات يظهر بوضوح عندما يزداد حجم العيّنة.

من ناحية ثانية، أظهر التحليل وجود أثر رئيس لكل متغيّر من المتغيّرات المستقلة الثلاثة (طول الاختبار، حجم العيّنة، شكل الاختبار) في المتغيّر التابع (الخطأ المعياري لتقدير معلمة التمييز)؛ فبالنسبة لمتغيّر الطول تبيّن أن قيم الخطأ المعياري لتقدير معلمة التمييز في الاختبارات الطويلة مهما كان شكلها (اختيار من متعدّد، صواب وخطأ) أقل منها مقارنة بالاختبارات القصيرة في العيّنة الصغيرة (250) فردًا، في حين تبيّن أن قيم الخطأ المعياري لتقدير معلمة التميز في الاختبارات القصيرة بصرف النظر عن شكل الاختبار (اختيار من متعدّد، صواب وخطأ) أقل منها مقارنة بالاختبارات الطويلة في العيّنة متوسطة الحجم (500) والعيّنة كبيرة الحجم (1015). أما بالنسبة لمتغيّر شكل الاختبار، فقد

تبيّن أن قيم الخطأ المعياري لتقدير معلمة التمييز في اختبارات الاختيار من متعدّد أقل منها إذا ما قورنت باختبارات الصواب والخطأ وبصرف النظر عن طول الاختبار أو حجم العيّنة، وبالنسبة لمتغيّر حجم العيّنة، فقد أظهر التحليل نوعًا من عدم الاتساق في النتائج، حيث تبيّن أن قيمة الخطأ المعياري لتقدير معلمة التمييز يقل كلما ازداد حجم العيّنة في اختبارات الصواب والخطأ بينما تبيّن العكس في اختبارات الاختيار من متعدّد.

ولمعرفة مصدر التباين الإجمالي الذي كشف عنه تحليل التباين للأثر الرئيس لمتغيّر حجم العيّنة (250، 500، 1015) فقد تمّ استخدام اختبار "بونفيروني" Bonferroni للمقارنات البعدية كما يتّضح من الجدول (25).

جدول (25)

نتائج اختبار "بونفيروني" Bonferroni للمقارنات البعدية للكشف عن مصدر التباين للأثر الرئيس لمتغيّر حجم العيّنة في قيمة الخطأ المعياري لتقدير معلمة التمييز

المتوسط الحسابي	حجم العيّنة	500	1015
0.248	250	0.013	0.031*
0.151	500		0.018*
0.133	1015		

* α = 0.01

يتّضح من الجدول (25) أعلاه أن مصدر الفروق الإجمالية التي كشف عنها تحليل التباين كان بسبب الاختلاف الجوهري بين متوسط قيم الخطأ المعياري لتقدير معلمة التمييز في العيّنة كبيرة الحجم (1015) عن كلٍ من متوسطي قيم الخطأ المعياري لتقدير معلمة التمييز في العيّنة متوسطة الحجم (500) وصغيرة الحجم (250).

السؤال الثاني: "ما الأثر الرئيس لكل متغيّر من متغيّرات طول الاختبار (طويل، قصير)، وحجم عيّنة المفحوصين (250، 500، 1015) مفحوصًا، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ)، وأثر التفاعلات الثنائية بين هذه المتغيّرات، والتفاعل الثلاثي بينها في دقة تقدير معلمة القدرة ؟

للإجابة عن هذا السؤال تمّ تقدير معلمة القدرة للمفحوصين في الاختبارات الأربعة (اختيار من متعدّد قصير، اختيار من متعدّد طويل، صواب وخطأ قصير، صواب وخطأ طويل) وذلك لكل مستوى من مستويات متغّير حجم العيّنة (250، 500، 1015) مفحوصًا على الترتيب كما هو موضّح في الجدول (26).

جدول (26)

المتوسطات الحسابية (م) والانحرافات المعيارية (ع) لتقديرات معلمة القدرة لكل اختبار من الاختبارات الأربعة موزّعة حسب مستويات متغيّر حجم العيّنة

حجم العيّنة	شكل الاختبار	طول الاختبار			
		طويل (50 فقرة)		قصير (25 فقرة)	
		م	ع	م	ع
250	اختيار من متعدّد	-0.0037	0.851	-0.0015	0.900
	صواب وخطأ	-0.0081	0.920	-0.0114	0.851
500	اختيار من متعدّد	-0.0105	0.929	0.0071	0.937
	صواب وخطأ	-0.0166	0.854	-0.0046	0.955
1015	اختيار من متعدّد	-0.0035	0.934	-0.0003	0.914
	صواب وخطأ	-0.0101	0.865	-0.0029	0.904

يُلاحظ من الجدول (26) أن قيم متوسطات تقديرات قدرة المفحوصين تقترب من الصفر بصرف النظر عن شكل الاختبار (اختيار من متعدّد، صواب وخطأ)، أو طول الاختبار (طويل، قصير)، أو حجم عيّنة المفحوصين (250، 500، 1015) مفحوصًا. كذلك الحال بالنسبة لقيم الانحراف المعياري

لتقديرات القدرة التي اقتربت من الواحد الصحيح بصرف النظر عن شكل الاختبار (اختيار من متعدّد، صواب وخطأ)، أو طول الاختبار (طويل، قصير)، أو حجم عيّنة المفحوصين (250، 500، 1015) مفحوصًا – أي أنه يمكن القول أنها تأخذ منحنى التوزيع السويَ – وتدل قيم متوسطات القدرة السالبة الى الضعف العام بقدرات المفحوصين الرياضية.

وللكشف عن دقة تقدير معلمة القدرة، والأثر الرئيس لكل متغيّر من المتغيّرات المستقلة (حجم العيّنة، شكل الاختبار، طول الاختبار)، وأثر التفاعلات الثنائية بينها، وأثر التفاعل الثلاثي في تقدير معلمة القدرة تمّ حساب الخطأ المعياري لتقدير معلمة القدرة، ثم تمّ حساب المتوسطات الحسابية والانحرافات المعيارية لقيم الخطأ المعياري لتقدير معلمة القدرة لكل اختبار من الاختبارات الأربعة (اختيار من متعدّد قصير، اختيار من متعدّد طويل، صواب وخطأ قصير، صواب وخطأ طويل) ولكل عيّنة من العيّنات الثلاث (250، 500، 1015) كما يوضّحها الجدول (27).

جدول (27)

المتوسطات الحسابية (م) والإنحرافات المعيارية (ع) لقيم الخطأ المعياري لتقدير معلمة القدرة موزّعة بحسب متغيرات حجم العيّنة وشكل الاختبار وطول الاختبار

شكل الاختبار	طول الاختبار	حجم العيّنة						المجموع	
		250		500		1015			
		م	ع	م	ع	م	ع	م	ع
اختيار من متعدّد	قصير	0.371	0.088	0.322	0.130	0.395	0.080	0.371	0.103
	طويل	0.412	0.042	0.318	0.103	0.334	0.093	0.341	0.095
صواب وخطأ	قصير	0.451	0.057	0.309	0.147	0.401	0.088	0.382	0.116
	طويل	0.285	0.136	0.410	0.066	0.436	0.058	0.407	0.091

يُلاحظ من الجدول (27) وجود فروق ظاهرية بين المتوسطات الحسابية لقيم الخطأ المعياري لتقدير معلمة القدرة في الاختبارات الأربعة (اختيار من متعدّد قصير، اختيار من متعدّد طويل، صواب وخطأ قصير، صواب وخطأ طويل) موزّعة حسب مستويات متغيّر حجم العيّنة (250، 500، 1015) مفحوصًا.

ولتعرّف دلالة الفروق بين قيم متوسطات الأخطاء المعيارية لتقدير معلمة القدرة باختلاف مستويات متغيّرات حجم عيّنة المفحوصين (250، 500، 1015) مفحوصًا، وطول الاختبار (طويل، قصير)، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ)، تمّ استخدام أسلوب تحليل التباين العاملي (2x2x3) ذو القياسات المتكرّرة repeated measures factorial ANOVA أحادي بين الأفراد وثنائي داخل الأفراد (One Between – Two Within Subjects Design)، حيث يمثّل متغيّر حجم العيّنة (250، 500، 1015) متغيّر بين الأفراد، بينما يمثّل متغيّري طول الاختبار (طويل، قصير)، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ) متغيّري داخل الأفراد كما هو موضّح في الجدول (28).

جدول (28)

نتائج تحليل التباين العاملي (2x2x3) ذو القياسات المتكرّرة للكشف عن أثر المتغيّرات المستقلة في تقدير معلمة القدرة

مصدر التباين	م. م	د.ح	م. م. م	قيمة "ف"	η^2 الجزئي
بين الأفراد					
حجم العيّنة (أ)	3.583	2	1.792	164.922**	0.158
داخل الأفراد					
شكل الاختبار (ب)	0.715	1	0.715	91.791**	0.050
طول الاختبار (ج)	0.102	1	0.102	12.775**	0.007
أ x ب	1.183	2	0.592	75.958**	0.079
أ x ج	2.296	2	1.148	144.443**	0.141
ب x ج	0.002	1	0.002	0.220	0.000
أ x ب x ج	4.987	2	2.494	308.576**	0.259

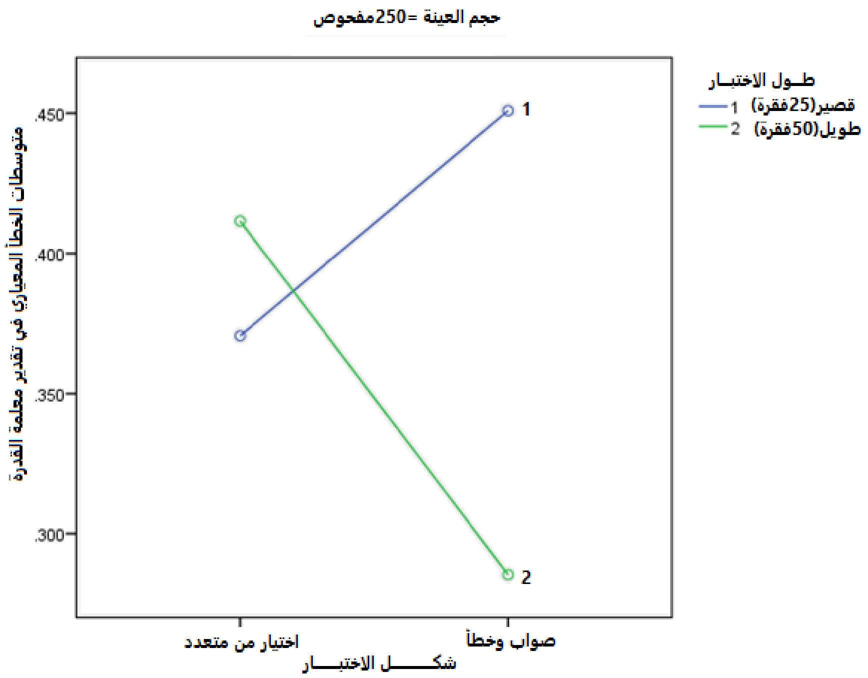
* $\alpha \leq 0.05$ (حيث تدل م.م : على مجموع المربعات، د.ح : درجات الحرية، م.م.م : معدل مجموع المربعات)

** $\alpha \leq 0.01$

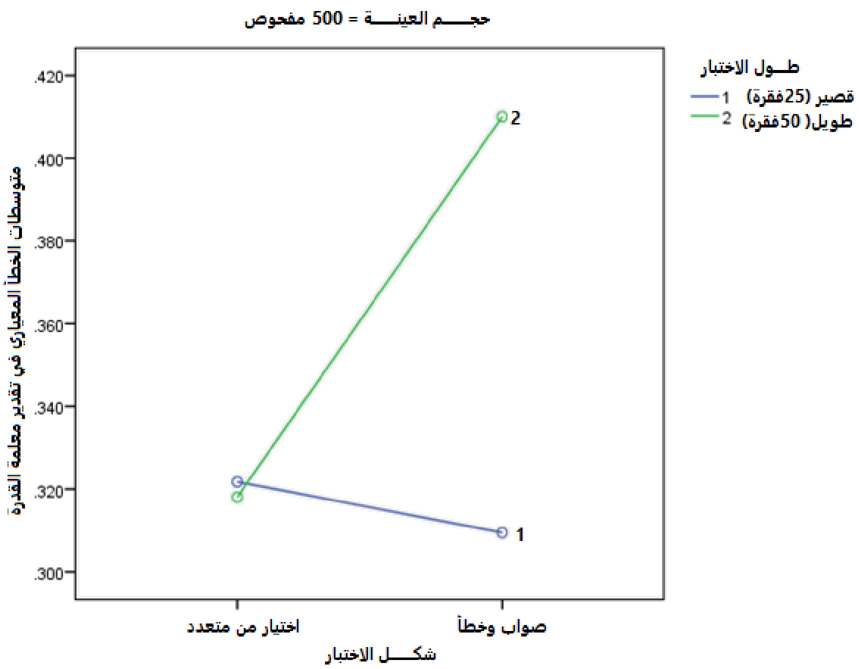
يتضح من الجدول (28) وجود أثر دال احصائياً للتفاعل الثلاثي بين كل من حجم العينة وشكل الاختبار وطول الاختبار مجتمعة في قيم الخطأ المعياري لتقدير معلمة القدرة وبالتالي على دقة التقدير.

كما يتّضح من الجدول أعلاه وجود أثر لجميع التفاعلات الثنائية المحتملة بين المتغيّرات المستقلة في المتغيّر التابع باستثناء تفاعل متغيّري طول الاختبار مع شكل الاختبار حيث لم يكن دال إحصائياً. ومن الملاحظ وجود أثر دال إحصائياً لكل متغير من المتغيرات الثلاثة كل على حدا، من جانبٍ آخر، كشف التحليل عن أن قيم حجم الأثر الجزئي (Partial η^2) لكل متغيّر من المتغيّرات المستقلة الثلاثة (حجم العيّنة، طول الاختبار، شكل الاختبار) تراوحت بين حجم الأثر الكبير والمعدوم؛

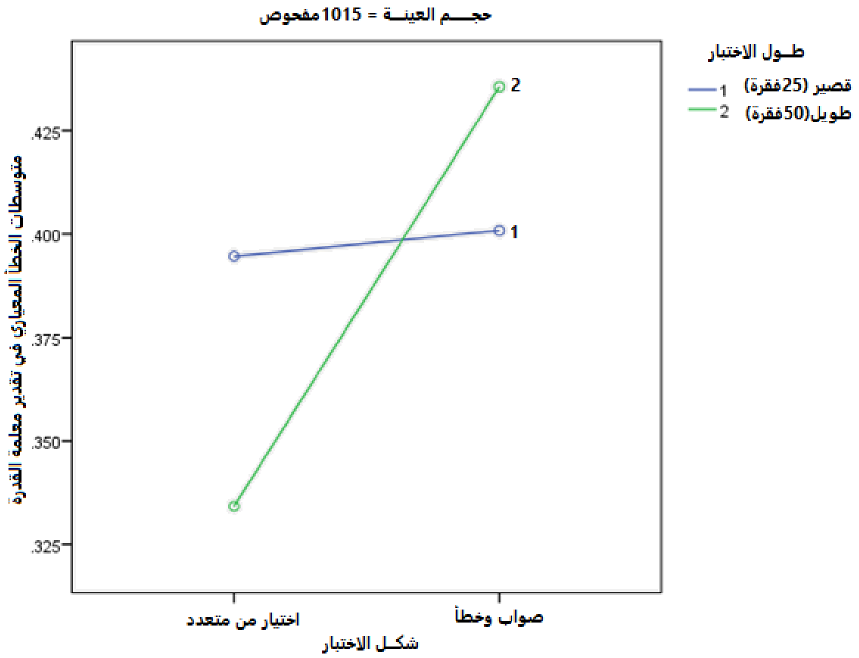
ففي حين أظهر التحليل وجود حجم أثر كبير يزيد عن (0.06) بحسب معيار "كوهن" Cohen لمتغيّر حجم العيّنة بلغ (15.8%)، تبيّن وجود أثر قليل لمتغيّر شكل الاختبار يقل عن (0.06)، وحجم أثر قليل جدًا أو معدوم يقل عن (0.01) لمتغيّر طول الاختبار بالرغم من وجود دلالة إحصائية لهذين المتغيّرين (شكل الاختبار، وطول الاختبار). وهذا يؤكد الاهتمام بالأهمية العملية (حجم الأثر) بصرف النظر عن الدلالة الإحصائية. أما بالنسبة للأهمية العملية للتفاعلات الثنائية بين المتغيرات الثلاثة مع بعضها بعضًا، فقد كشف التحليل عن حجم أثر كبير يزيد عن (0.13) لتفاعل متغيّري حجم العيّنة وطول الاختبار، بينما كان حجم الأثر لتفاعل متغيّري حجم العيّنة وشكل الاختبار متوسطًا حيث بلغ (7.90%)، في حين لم يتبيّن وجود أي أثر من الناحيتين الإحصائية والعملية لتفاعل متغيّري طول الاختبار وشكل الاختبار. وبالنسبة للتفاعل الثلاثي بين العوامل الثلاثة، فقد بلغ حجم الأثر (0.259)، وهو حجم أثر كبير بحسب معيار "كوهن" Cohen. وهذا يعني أن التفاعل الثلاثي بين المتغيّرات المستقلة في التصميم (حجم العيّنة، شكل الاختبار، طول الاختبار) يفسّر ما مقداره (25.90%) من التباين في تقديرات الخطأ المعياري لمعلمة القدرة. وتوضّح الأشكال (10، 11، 12) أثر التفاعلات التي كشف عنها التحليل.



شكل (10): التفاعل الثنائي بين متغيّري شكل الاختبار وطول الاختبار عند العينة الصغيرة (250 مفحوصاً) في قيمة الخطأ المعياري لتقدير معلمة القدرة



شكل (11): التفاعل الثنائي بين متغيّري شكل الاختبار وطول الاختبار عند العينة المتوسطة (500 مفحوص) في قيمة الخطأ المعياري لتقدير معلمة القدرة



شكل (12): التفاعل الثنائي بين متغيّري شكل الاختبار وطول الاختبار عند العينة الكبيرة (1015 مفحوصاً) في قيمة الخطأ المعياري لتقدير معلمة القدرة

يتبيّن من الأشكال (10، 11، 12) أن قيم الخطأ المعياري في تقدير معلمة القدرة تنخفض في اختبارات الاختيار من متعدّد الطويلة مقارنة باختبارات الصواب والخطأ عند العينة متوسطة الحجم (500 مفحوص) والعينة كبيرة الحجم (1015 مفحوص)، أما عند العينة صغيرة الحجم (250 مفحوص) فإن قيم الخطأ المعياري تنخفض في اختبارات الصواب والخطأ لتحقق أقل متوسط في الخطأ المعياري على الاطلاق، مما يعني دقة تقدير أعلى، أما في الاختبارات القصيرة فإن قيم الخطأ المعياري في تقدير معلمة القدرة تبدي انخفاضاً في اختبارات الاختيار من متعدد عند العينتين الصغيرة (250 مفحوص) والكبيرة (1015 مفحوص)، و عند العينة المتوسطة (500 مفحوص) كانت قيم الخطأ أقل عند اختبار الصواب والخطأ. إلا أن الفرق بين قيم الخطأ المعياري في هذه الاختبارات يظهر بوضوح عندما يزداد حجم العيّنة.

ولمعرفة مصدر التباين الإجمالي الذي كشف عنه تحليل التباين للأثر الرئيس لمتغيّر حجم العيّنة (250، 500، 1015) فقد تمّ استخدام اختبار "بونفيروني" Bonferroni للمقارنات البعدية كما يتّضح من الجدول (29).

جدول (29)

نتائج اختبار "بونفيروني" Bonferroni للمقارنات البعدية للكشف عن مصدر التباين للأثر الرئيس لمتغيّر حجم العيّنة في قيمة الخطأ المعياري لتقدير معلمة القدرة

المتوسط الحسابي	حجم العيّنة	500	1015
0.380	250	*0.040	*0.012
0.340	500		*0.051
0.391	1015		

* α = 0.05

يتّضح من الجدول (29) أن متوسط قيم الخطأ المعياري لتقدير معلمة القدرة في العيّنة الأكبر حجمًا (1015) كان هو الأكبر، يليه في العيّنة الأصغر حجمًا (250)، وأخيرًا في العيّنة متوسطة الحجم (500). ويتبيّن من الجدول أيضاً أن الفروق الثلاثة بين متوسطات قيم الخطأ المعياري لتقدير معلمة القدرة في العيّنات الثلاث ذات دلالة إحصائية عند مستوى دلالة يقل عن (0.05)؛ أي أن هناك فروق جوهرية ذات دلالة إحصائية (α = 0.05) بين المتوسطات الثلاثة لقيم الخطأ المعياري لتقدير معلمة القدرة.

السؤال الثالث: "ما الأثر الرئيس لكل متغيّر من متغيّرات حجم عيّنة المفحوصين (250، 500، 1015) مفحوصًا، وطول الاختبار (طويل، قصير)، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ)، وأثر التفاعلات الثنائية بين هذه المتغيّرات، والتفاعل الثلاثي بينها في دقة تقدير دالة معلومات الاختبار ؟

للإجابة عن هذا السؤال تمّ حساب دالة المعلومات في الاختبارات الأربعة (اختيار من متعدّد قصير، اختيار من متعدّد طويل، صواب وخطأ قصير، صواب وخطأ طويل) وذلك لكل مستوى من مستويات متغّير حجم العيّنة (250، 500، 1015) مفحوصًا على الترتيب على متّصل القدرة الذي يمتد من (3−) لوجيت إلى (3) لوجيت كما هو موضّح في الجدول (30).

جدول (30)

قيم دالة المعلومات المقدرة للاختبارات المختلفة عند مستويات قدرة مختلفة

حجم العينة	نوع الاختبار	عدد الفقرات														متوسط دالة المعلومات
		50 فقرة							25 فقرة							
		مستوى القدرة							مستوى القدرة							
		−3	−2	−1	صفر	1	2	3	−3	−2	−1	صفر	1	2	3	
250	اختيار من متعدّد	3.48	4.94	5.86	5.83	5.18	4.20	3.18	1.14	2.47	4.97	7.60	7.34	4.97	2.75	**4.57**
	صواب وخطأ	2.84	4.42	6.39	9.67	15.9	11.3	4.92	1.69	3.35	5.51	5.43	3.91	2.70	1.83	**5.70**
500	اختيار من متعدّد	2.60	4.51	7.28	9.78	9.79	7.35	4.64	0.87	2.34	5.88	10.8	10.8	6.26	2.94	**6.13**
	صواب وخطأ	3.90	5.63	6.45	6.06	4.92	3.56	2.39	0.74	2.27	6.84	14.3	13.0	6.08	2.69	**5.63**
1015	اختيار من متعدّد	2.64	4.47	6.92	8.88	8.72	6.67	4.36	1.16	2.50	4.89	7.04	6.50	4.30	2.45	**5.11**
	صواب وخطأ	3.61	4.91	5.54	5.27	4.35	3.20	2.22	1.57	3.31	6.00	7.81	5.89	3.27	1.82	**4.20**
المتوسط		**3.18**	**4.81**	**6.41**	**7.58**	**8.14**	**6.05**	**3.62**	**1.20**	**2.71**	**5.68**	**8.83**	**7.91**	**4.60**	**2.41**	

يُلاحظ من الجدول (30) أن قيمة دالة المعلومات في اختبارات الاختيار من متعدّد تزداد كلما اقترب مستوى القدرة من منتصف المقياس؛ أي عندما يكون مستوى القدرة مساويًا للصفر على متّصل القدرة وبصرف النظر عن حجم العيّنة وطول الاختبار. وبالمثل في اختبارات الصواب والخطأ، فيُلاحظ أن قيمة دالة المعلومات تزداد كلما اقترب مستوى القدرة من منتصف المقياس وبصرف النظر عن حجم العيّنة وطول الاختبار باستثناء الاختبار الطويل (50 فقرة) عندما كان حجم عيّنة المفحوصين (250) مفحوصًا حيث كانت أكبر قيمة لدالة المعلومات عندما كان مستوى القدرة (1) على متّصل القدرة. ويُلاحظ أيضاً أن أقصى قيمة لمتوسط قيم دالة المعلومات للاختبار الطويل وبصرف النظر عن نوع الاختبار وعن حجم العيّنة كانت عند مستوى القدرة (1)، مع أن هذه القيمة تقترب من متوسط قيم دالة المعلومات للاختبار الطويل وعن حجم العيّنة وبصرف النظر عن نوع الاختبار عند مستوى القدرة (صفر). أما في الاختبار القصير، فقد تبيّن أن متوسط قيم دالة المعلومات يبلغ ذروته عند مستوى القدرة (صفر) وبصرف النظر عن نوع الاختبار وعن حجم عيّنة المفحوصين. من ناحية ثانية، يتبيّن من الجدول نفسه أن قيمة متوسط دالة المعلومات تكون أكبر ما يمكن في اختبار الاختيار من متعدّد بصرف النظر عن طوله عندما يكون حجم العيّنة (500). وتبيّن أيضاً أنه كلما ازداد حجم العيّنة تكون قيمة دالة المعلومات في اختبار الاختيار من متعدّد أكبر منها مقارنة باختبار الصواب والخطأ، ويُستثنى من ذلك الحالة التي يكون فيها حجم العيّنة (250) مفحوصًا. يُستنتج مما سبق وجود مؤشرات على حُسن تدريج الاختبارات بشكل عام سواء كانت من نوع الاختيار من متعدّد أو من نوع الصواب والخطأ، وسواء كانت من النوع الطويل (50 فقرة) أو من النوع القصير (25) فقرة، ومهما كان حجم العيّنة.

ولتسهيل المقارنة بين الاختبارات تم حساب الكفاءة النسبية لها اعتمادًا على العلاقة الآتية، ويُلخّص الجدول (31) الآتي هذه المقارنات.

$$RE(\theta) = \frac{I_A(\theta)}{I_B(\theta)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots [19] \quad -$$

حيث $RE(\theta)$: الفعالية النسبية (أو الكفاءة النسبية).

$I_A(\theta)$: دالة المعلومات للإختبار A

$I_B(\theta)$: دالة المعلومات للإختبار B

وتكون الدالتان معرفتين على متصل قدرة مشترك (θ).

جدول (31)

قيم الكفاءة النسبية لدالة معلومات الاختبار وفقاً لحجم العينة وشكل الاختبار وطول الاختبار

المقارنة وحجم العيّنة	50 فقرة			25 فقرة		
	مستوى القدرة			مستوى القدرة		
	−3	0	3	−3	0	3
اختيار من متعدد (250) مع صواب وخطأ (250)	1.23	0.60	0.64	0.67	1.40	1.50
اختيار من متعدد (250) مع اختيار من متعدد (500)	1.34	0.60	0.68	1.31	0.71	0.93
اختيار من متعدد (250) مع صواب وخطأ (500)	0.89	0.96	1.33	1.55	0.53	1.02
اختيار من متعدد (250) مع اختيار من متعدد (1015)	1.32	0.66	0.73	0.98	1.08	1.12
اختيار من متعدد (250) مع صواب وخطأ (1015)	0.96	1.11	1.43	0.73	0.97	1.51
صواب وخطأ (250) مع اختيار من متعدد (500)	1.09	0.99	1.06	1.95	0.50	0.62
صواب وخطأ (250) مع صواب وخطأ (500)	0.73	1.60	2.06	2.30	0.38	0.68
صواب وخطأ (250) مع اختيار من متعدد (1015)	1.07	1.09	1.13	1.45	0.77	0.75
صواب وخطأ (250) مع صواب وخطأ (1015)	0.78	1.84	2.09	1.08	0.70	1.01
اختيار من متعدد (500) مع صواب وخطأ (500)	0.67	1.61	1.94	1.18	0.75	1.09
اختيار من متعدد (500) مع اختيار من متعدد (1015)	0.99	1.10	1.06	0.75	1.53	1.20
اختيار من متعدد (500) مع صواب وخطأ (1015)	0.72	1.86	1.08	0.55	1.38	1.62
صواب وخطأ (500) مع اختيار من متعدد (1015)	1.48	0.68	0.50	0.63	2.04	1.10
صواب وخطأ (500) مع صواب وخطأ (1015)	1.08	1.15	1.08	0.47	1.84	1.48
اختيار من متعدد (1015) مع صواب وخطأ (1015)	0.73	1.68	1.97	0.74	0.90	1.35

واعتماداً على ما أشار إليه كل من (Crocker & Algina, 1986) أنه لأي درجة سمة كامنة عندما تكون (RE(θ,A,B)>1) فإن A يكون قياس أفضل (أي عندما تكون الكفاءة النسبية أكبر من واحد فإن الاختبار الأول يكون أفضل من الاختبار الثاني)، وبصورة مشابهة عندما تكون درجة السمة الكامنة (RE(θ,A,B)<1) فإن *B* يكون قياس أفضل للسمة الكامنة (أي عندما تكون الكفاءة النسبية أقل من واحد فإن الاختبار الثاني يكون أفضل من الاختبار الأول)؛ وذلك بالاعتماد على العلاقة [19] أعلاه.

يُلاحظ من الجدول (31) أن كفاءة الاختبار لا تتأثر بمتغيّر طول الاختبار عند مستويات القدرة المختلفة، وأن كفاءة الاختبار سواء كان من نوع الاختيار من متعدّد أو الصواب والخطا تتناسب تناسبًا عكسيًا مع حجم العينة، كما أن اختبارات الاختيار من متعدّد تتفوق على اختبارات الصواب والخطأ في حال تساوي حجم العيّنة. أما في حال اختلاف حجم العيّنة، فقد تناوبت كفاءة الاختبارات من النوعين؛ إذ تفوقت اختبارات الاختيار من متعدّد عندما كان حجم العيّنة في هذه الاختبارات (250) أو (500) على اختبارات الصواب والخطأ عندما كان حجم العيّنة في هذه الاختبارات (500) أو (1015)، بينما تفوقت اختبارات الصواب والخطأ عندما كان حجم العيّنة في هذه الاختبارات (250) أو (500) على اختبارات الاختيار من متعدّد عندما كان حجم العيّنة في هذه الاختبارات (500) أو (1015).

من ناحية ثانية، تمّ حساب قيم الخطأ المعياري المصاحبة لعملية تقدير دالة معلومات الاختبار عند مستويات مختلفة من القدرة كما يتّضح من الجدول (32).

جدول (32)

قيم الخطأ المعياري (SEE) في تقدير دالة معلومات كلٍ من الاختبارات الأربعة عند مستويات قدرة مختلفة

حجم العينة	طول الاختبار	طويل (50 فقرة)							قصير (25 فقرة)						
	شكل الاختبار	مستويات القدرة							مستويات القدرة						
		−3	**−2**	**−1**	**0**	**1**	**2**	**3**	**−3**	**−2**	**−1**	**0**	**1**	**2**	**3**
250	اختيار من متعدد	.536	.450	.413	.414	.439	.488	.561	.938	.636	.448	.363	.370	.448	.604
	صواب وخطأ	.594	.476	.396	.322	.251	.298	.451	.769	.546	.426	.429	.506	.608	.740
500	اختيار من متعدد	.620	.471	.371	.320	.320	.369	.464	1.074	.653	.413	.305	.304	.400	.584
	صواب وخطأ	.507	.422	.394	.406	.451	.530	.646	1.170	.664	.382	.264	.277	.406	.610
1015	اختيار من متعدد	.61	.473	.380	.336	.339	.387	.479	.927	.633	.452	.377	.392	.482	.640
	صواب وخطأ	.52	.451	.425	.436	.480	.558	.672	.800	.550	.408	.358	.412	.553	.247

يتّضح من الجدول (32) أن قيم الخطأ المعياري في الاختبارات الطويلة (50 فقرة) تراوحت بين (0.251) كحد أدنى عند اختبار الصواب والخطأ للعينة ذات الحجم (250) مفحوصًا وعند مستوى القدرة (1)، و (0.672) كحد أعلى عند اختبار الصواب والخطأ للعينة ذات الحجم (1015) مفحوصًا وعند مستوى القدرة (3). أما في الاختبارات القصيرة (25 فقرة)، فقد تراوحت قيم الخطأ المعياري بين (0.264) كحد أدنى لاختبار الصواب والخطأ للعينة المكوَنة من (500) مفحوصًا عند مستوى القدرة (0)، و (1.17) كحد أعلى لاختبار الصواب والخطأ للعينة المكوَنة من (500) مفحوصًا، عند مستوى القدرة المنخفض (−3).

وللكشف عن الآثار الرئيسة لمتغيّرات طول الاختبار (قصير، طويل)، وشكل الاختبار (اختيار من متعدّد، صواب وخطأ)، وحجم عيّنة المفحوصين (250، 500، 1015)، بالإضافة إلى أثر التفاعلات الثنائية الممكنة بين هذه المتغيّرات، وأثر التفاعل الثلاثي بينها في قيم الخطا المعياري لتقدير دالة المعلومات تمّ استخدام أسلوب تحليل التباين العاملي (2x2x3) ذو القياسات المتكرّرة repeated measures factorial ANOVA كما هو موضّح في الجدول (33).

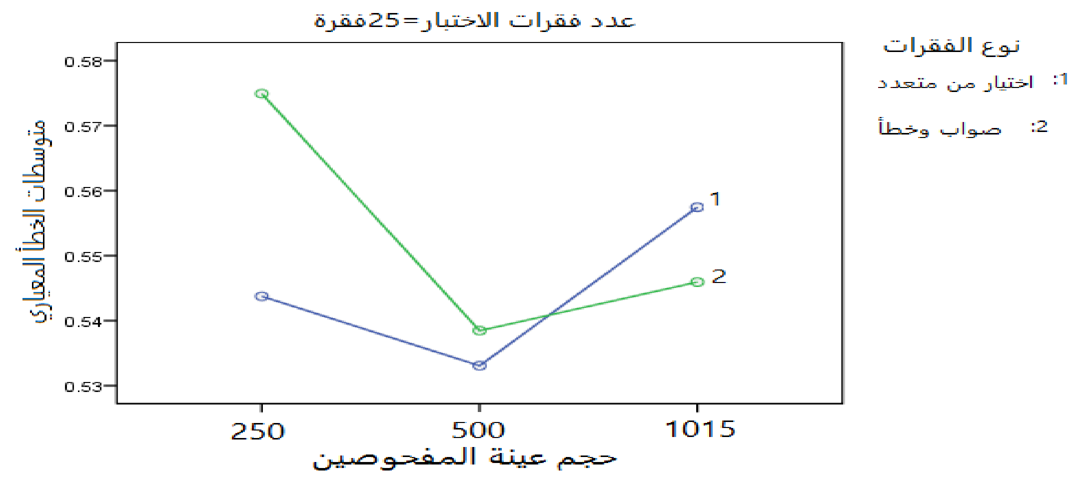
جدول (33)

نتائج التحليل العاملي للتباين (2x2x3) ذو القياسات المتكرّرة للكشف عن أثر المتغيرات المستقلة في تقدير دالة معلومات الاختبار عند جميع مستويات القدرة

مصدر التباين	م. م	د.ح	م. م. م	قيمة "ف"	η^2 الجزئي
بين الأفراد					
حجم العيّنة (أ)	0.005	2	0.002	0.078	0.000
داخل الأفراد					
شكل الاختبار (ب)	0.005	1	0.005	0.154	0.000
طول الاختبار (ج)	0.202	1	0.202	6.803*	0.068
أ x ب	0.014	2	0.007	0.228	0.005
أ x ج	0.007	2	0.004	0.124	0.009
ب x ج	0.001	1	0.001	0.029	0.003
أ x ب x ج	0.037	2	0.019	0.627	0.028

* $\alpha \leq 0.05$ (حيث تدل م.م : على مجموع المربعات، د.ح : درجات الحرية، م.م.م : معدل مجموع المربعات)

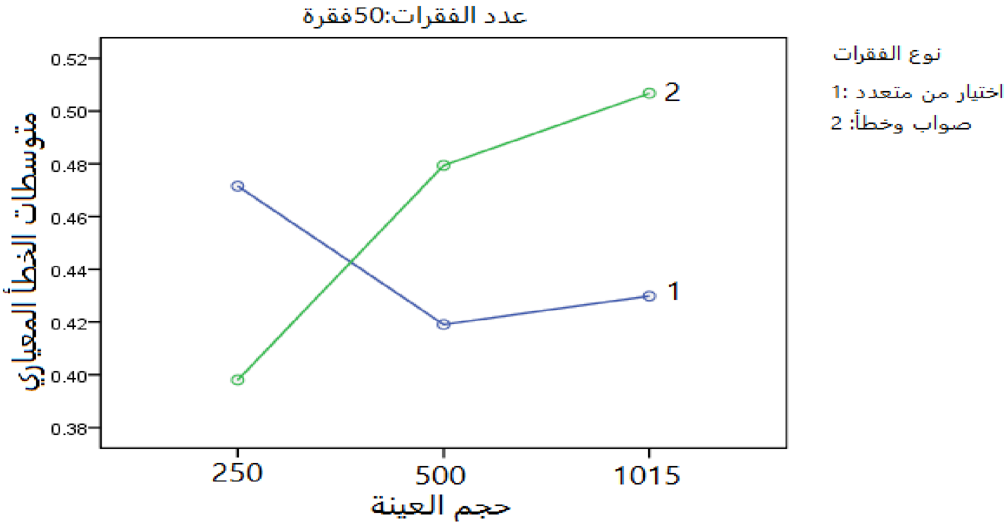
يُلاحظ من الجدول (33) عدم وجود أثر ذو دلالة إحصائية ($\alpha = 0.05$) للمتغيرات المستقلة، والتفاعلات الثنائية والثلاثية المختلفة بينها في تقدير دالة معلومات الاختبار باستثناء متغيّر طول الاختبار كما يتّضح من الأشكال (13، 14).



شكل (13): تفاعل حجم العينة وشكل الاختبار في الخطأ المعياري لتقدير دالة المعلومات للاختبار القصير

يتّضح من الشكل (13) ارتفاع في قيم الخطأ المعياري في تقدير دالة معلومات اختبارات الصواب والخطأ بشكل عام، وانخفاض في قيم الخطأ المعياري لاختبارات الاختيار من متعدّد في العيّنات صغيرة ومتوسطة الحجم (250، 500) مفحوصًا. أما عندما يزداد حجم العيّنة تصبح قيم الخطأ المعياري لاختبارات الصواب والخطأ أقل منها لاختبارات الاختيار من متعدّد؛ مما يعني أن دقة التقدير كانت أكبر لاختبارات الاختيار من متعدّد عند العينتين الأقل حجمًا، بينما أصبحت دقة التقدير أكبر لاختبارات الصواب والخطأ عند العينة الأكبر حجمًا (1015) مفحوصًا، وكانت أقل قيمة للخطأ المعياري في تقدير دالة معلومات الاختبار في العيّنة متوسطة الحجم (500) مفحوصًا لاختبار الاختيار من

متعدد مما يعني دقة تقدير أكبر، هذا كله عندما يتكون الاختبار من 25 فقرة. أما عندما يتكون الاختبار من (50 فقرة) فيوضّحه شكل (14) الآتي:



شكل (14): تفاعل حجم العينة وشكل الاختبار في قيم الخطأ المعياري لتقدير دالة المعلومات للاختبار الطويل

يُلاحظ من الشكل (14) أن قيم متوسطات الخطأ المعياري لاختبار الصواب والخطأ أقل منها لاختبار الاختيار من متعدد في العيّنة صغيرة الحجم (250) مفحوصًا مما يؤشر على دقة أعلى في التقدير، بينما تقل هذه القيم في اختبارات الاختيار من متعدّد مقارنة باختبارات الصواب والخطأ عند زيادة حجم العينة إلى (500، 1015) مفحوصًا مما يدل على دقة أكبر في التقدير عند اختبارات الاختيار من متعدّد عند الحجمين الأكبر للعينة مع ملاحظة دقة تقدير أكبر عند حجم (500) مفحوصًا عنها عند حجم (1015) مفحوصًا لتدني متوسط الخطأ المعياري في تقدير دالة معلومات الاختبار عند (500) مفحوصًا.

وعند مقارنة قيم متوسطات الخطأ المعياري في الشكل (13) والشكل (14) يتبيّن تدني في هذه القيم بشكل عام عند الاختبارات الطويلة (50 فقرة) مما يعني دقة أكبر في تقدير دالة المعلومات، وهذا

ما يؤكده تحليل التباين العاملي في جدول (33) من وجود دلالة إحصائية لأثر طول الاختبار في متوسطات الخطأ المعياري، وبالتالي في دقة التقدير لصالح الاختبارات الطويلة.

من جانبٍ آخر، كشف التحليل عن أن قيم حجم الأثر الجزئي (Partial η^2) لكل متغيّر من المتغيّرات المستقلة الثلاثة (حجم العيّنة، طول الاختبار، شكل الاختبار) تراوحت بين حجم الأثر الكبير والمعدوم؛ ففي حين أظهر التحليل وجود حجم أثر كبير يزيد عن (0.06) بحسب معيار "كوهن" Cohen لمتغيّر طول الاختبار بلغ (6.8%)، تبيّن عدم وجود أثر لكل من المتغيّرين شكل الاختبار وحجم العينة، وحجم أثر قليل جدًا أو معدوم يقل عن (0.01) لتفاعل كل من المتغيرين (حجم العينة، شكل الاختبار)، وكذلك لتفاعل (حجم العينة، وطول الاختبار)، وتفاعل (شكل الاختبار، وطول الاختبار. أما بالنسبة للأهمية العملية لتفاعل العوامل الثلاثة مع بعضها بعضًا، فقد كشف التحليل عن حجم أثر قليل يزيد عن (0.01) بقليل ، فقد بلغ حجم الأثر (0.028)، وهو حجم أثر قليل بحسب معيار "كوهن" Cohen. وهذا يعني أن التفاعل الثلاثي بين المتغيّرات المستقلة في التصميم (حجم العيّنة، شكل الاختبار، طول الاختبار) يفسّر ما مقداره (2.8%) فقط من التباين في تقديرات الخطأ المعياري دالة معلومات الاختبار.

الفصل الخامس

مناقشة النتائج والتوصيات

حاولت هذه الدراسة الكشف عن أثر ثلاثة من المتغيرات المستقلة وهي حجم العينة، طول الاختبار، شكل الاختبار (أو نوع الفقرات)، والتفاعل بين هذه المتغيرات في دقة تقدير كل من معالم الفقرات، وقدرة الأفراد، ودالة معلومات الاختبار. ولهذا الغرض، استخدم الباحث في هذه الدراسة اختبار تحصيلي في مادة الرياضيات للصف الثامن الأساسي للفصل الدراسي الأول. وتألفت عينة الدراسة من طلبة المدارس التابعة لمديرية قصبة عمان (عمان الأولى) – سابقاً – وقُسمت عينة الدراسة إلى ثلاث مجموعات هي (250، 500، 1015) طالبًا وطالبة لتشكل عينة صغيرة نسبياً ومتوسطة وكبيرة نسبياً. وبالنسبة لطول الاختبار والمتمثل بعدد فقراته، فقد تم بناء اختبارين اثنين؛ تكوّن الأول من (50) فقرة، بينما تكوّن الثاني من (25) فقرة، يمثل الأول اختبار طويل والثاني اختبار قصير بالنسبة لمستوى هذه الفئة من الطلبة وبالنظر إلى قدراتهم الجسمية والعقلية. أما بالنسبة لشكل الاختبار فقد تم اعتماد نوعين شائعين ومعروفين للطلبة يسهل التعامل معهما، يتمثل النوع الأول بفقرات من نوع الاختيار من متعدّد بأربعة بدائل تم وضعها بعناية، بينما كان الاختبار الثاني من نوع الصواب والخطأ. وقد تمّ التحقق من صدق وثبات الاختبارات الأربعة (اختيار من متعدّد قصير، اختيار من متعدّد طويل، صواب وخطأ قصير، صواب وخطأ طويل) توفّر لها دلالات صدق وثبات مقبولة لأغراض هذه الدراسة.

ونظراً لتطبيق الاختبار على طلبة العينة والحصول على بيانات حقيقية ومباشرة من الميدان فقد شاب هذه البيانات ملوَثات – والتي تتمثل أسبابها في جدّية الطلبة في الإجابة، وتفاوت اهتمامهم بالاختبار، وتنوع هذه المؤثرات واختلاف الظروف المحيطة عند تطبيق كل اختبار، وكذا الفوارق الطبيعية

بين الإناث والذكور بالحرص والاهتمام، إضافة لتمايز المعلمين في تعاونهم المنشود – فقد أدى كل ذلك وغيره إلى عدم تحقّق بعض الافتراضات المطلوبة لإتمام عملية تحليل التباين بالشكل المأمول.

لذلك عند مطابقة الفقرات والأفراد للنموذج اللوجستي ثلاثي المعلمة فقد تبيّن، عدم مطابقة الكثير من الفقرات ذلك النموذج مما يستدعي حذفها، والتأثير سلباً على نتائج الدراسة. وعند إجراء المطابقة للنموذج ثنائي المعلمة أظهرت النتائج مطابقة معظم الفقرات وجميع الأفراد لهذا النموذج ذلك بالاعتماد على نتائج قيم مربع كاي (χ^2) ودلالتها الاحصائية حيث تكون الفقرة غير مطابقة للنموذج إذا كانت احتمالية قيمة (χ^2) أقل من (0.05)، لذا اقتصرت معالم الفقرات التي جرى تقديرها على معلمتين هما (الصعوبة والتمييز) – للأسف– دون معلمة التخمين. مع ضرورة وجود معلمة التخمين نظراً كون الاختبارات في هذه الدراسة تحتمل أن يلجأ المفحوص الى التخمين للوصول الى الإجابة الصحيحة. من جانب آخر، لم يتحقق افتراض سويَة توزيع البيانات بالنظر إلى المتوسط الحسابي والانحراف المعياري، وكان منحنى بعض الاختبارات يأخذ التواءاً سالباً والبعض الآخر يأخذ التواءاً موجباً.

وبالنظر إلى النتائج التي أسفرت عنها الدراسة، والرجوع إلى الجداول والأشكال البيانية التوضيحية الواردة في الفصل الرابع، وربط النتائج بنتائج الدراسات السابقة ذات الصلة – التي استطاع الباحث الوصول اليها – الواردة في الفصل الثاني يمكن استخلاص ما يأتي:

للإجابة عن السؤال الأول والمتعلق بأثر حجم العينة على دقة تقدير معالم الفقرة ، فبالنسبة لمعلمة الصعوبة فإن أقل قيمة لمتوسط الخطأ المعياري وبالتالي دقة أعلى في التقدير عندما كان حجم عينة المفحوصين (1015) مفحوصًا، ذلك عند غض النظر عن طول الاختبار وشكل الاختبار، ويلي ذلك في دقة التقدير عندما كان عدد المفحوصين (500) مفحوصًا، وهذا يتفق مع الكثير من الدراسات المتعلقة بهذا الشأن مثل دراسة ليم وهون وهاردويل وميشيل (1997)، وتتفق مع ما توصل اليه الثوابية

(2010)، وتتفق أيضاً مع ما توصل اليه كل من كينج بينج وتشيرس بوجود أثر لحجم العينة على دقة تقدير معالم الفقرة. وتتفق مع نتائج دراسة عبابنة (2004) التي توصلت الى أن الدقة في تقدير معالم الفقرة تزداد بزيادة حجم العينة، حيث تبيّن في الدراسة الحاليّة أن أقل دقة في تقدير الصعوبة كانت عندما كان حجم العينة (250) مفحوصًا. وفيما يتعلق بتقدير معلمة التمييز كان أقل متوسط للخطأ المعياري عند حجم العينة الأصغر (250 مفحوص)، مما يؤشر على دقة أعلى في تقدير تمييز الفقرات، وبشكل عام تبيّن وجود فروق جوهرية ذات دلالة احصائية عند مستوى الدلالة (α=0.05) بين متوسطات الأخطاء المعيارية في تقدير معالم الفقرة (الصعوبة، والتمييز) تُعزى لمتغير حجم العيّنة.

وفيما يتعلق بأثر طول الاختبار (عدد فقرات الاختبار) في تقدير معلمة صعوبة الفقرات، وبصرف النظر عن حجم العينة ونوع الفقرات، كان متوسط الخطأ المعياري أقل عندما كان عدد فقرات الاختبار (50 فقرة) مما يدل على دقة أعلى في التقدير. وبالنسبة لمعلمة التمييز، فقد كان متوسط الخطأ المعياري الأقل في الاختبار الطويل (50 فقرة)، بالتالي دقة تقدير أعلى مقارنة بالاختبار القصير(25 فقرة).

فيما يتعلق بأثر نوع الفقرة (أو شكل الفقرة) على دقة التقدير، تشير النتائج الى أن دقة تقدير كل من معلمة الصعوبة والتمييز كانت أكبر عندما كانت فقرات الاختبار من نوع الاختيار من متعدد، بالتالي كان تقدير كل من الصعوبة والتمييز أقل دقة للأختبارات ذات فقرات من نوع الصواب والخطأ؛ ذلك اعتماداً على قيم متوسطات الخطأ المعياري في تقدير المعالم، وبغض التظر عن المتغيرات الأخرى، وتتفق هذه النتائج مع نتائج دراسة العنزي (2010) ودراسة المصري (2009) فيما يتعلق بتقدير معلمتي الصعوبة والتمييز، حيث توصلا الى وجود فروق ذات دلالة احصائية عند مستوى (α =0.05) في تقدير صعوبة الفقرات وتمييزها تُعزى الى شكل فقرات الاختبار.

ويمكن القول أن أثر تفاعل حجم العينة وطول الاختبار وشكل الاختبار على دقة تقدير معلمة صعوبة الفقرات كان لصالح تفاعل حجم العينة الأكبر والاختبار الأطول المكوَّن من فقرات الاختيار من متعدد، وتتوافق هذه النتيجة مع ما توصل اليه الحجيلي (2012) بأن دقة تقدير الصعوبة تختلف بشكل دال عند طول الاختبار وحجم العينة والتفاعل بينهما، وتتفق أيضاً مع ما ذهب اليه الدرابيع (2001).

بالنسبة لتقدير معلمة التمييز فقد كانت دقة التقدير الأعلى لها عند تفاعل حجم العينة الأصغر والاختبار الأطول المكوَنه فقراته من الاختيار من متعدد أيضاً. وهذا يتفق مع نتائج الكثير من الدراسات السابقة المشار اليها عدا ما يتعلق بحجم العينة.

للإجابة على سؤال الدراسة الثاني فيما يتعلق بتقدير معلمة القدرة، فقد كانت متوسطات الخطأ المعياري في التقدير متقاربة بشكل كبير عند المستويات الثلاثة لحجم العيّنة مما يدل على تقارب في دقة التقدير، وكان متوسط الخطأ المعياري لتقدير معلمة القدرة في العيّنة صغيرة الحجم (250) مفحوصًا أقل منه مقارنة بكلٍ من العيّنة متوسطة الحجم (500) مفحوصًا، والعيّنة كبيرة الحجم (1015) مفحوصًا مما يؤشر على دقة أعلى في تقدير القدرة في العيّنات (250، 500، 1015) على الترتيب وبصرف النظر عن طول الاختبار وشكله، وهذه النتيجة لا تتفق مع ما توصلت إليه بعض الدراسات. وبالنسبة لأثر طول الاختبار في دقة تقدير معلمة القدرة فقد كان أقل متوسط حسابي للخطأ المعياري عند الاختبار الطويل (50 فقرة)، أي أن دقة التقدير أعلى. وهذا يتفق مع ما توصل إاليه جلاس (2005) الذي أشار إلى أنه بزيادة عدد فقرات الاختبار تقل الأخطاء المعيارية وبالتالي تزداد دقة تقدير معلمة القدرة، وتنسجم هذه النتيجة أيضاً مع نتائج دراسة عبابنة (2004) و كل من الدرابيع (2001) وبارنس ووايز (1991).

وفيما يتعلق بأثر نوع الفقرات على دقة تقدير معلمة القدرة أشارت النتائج الى أن دقة التقدير الأعلى تتحقق عند استخدام فقرات على شكل صواب وخطأ، حيث حققت أدنى قيمة في متوسطات

الخطأ المعياري في التقدير، وقد يعود ذلك – في نظر الباحث– الى زيادة احتمال وصول الطالب الى الاجابة الصحيحة مقارنة مع فقرات الاختيار من متعدد بالاضافة الى زيادة التخمين في الوصول الى الاجابة. واستناداً على ما سبق يمكن القول أن تفاعل كل من حجم العينة الأصغر والاختبار الأطول الذي فقراته على شكل صواب وخطأ كان له الأثر الأكبر على دقة تقدير معلمة القدرة مقارنة مع حدود وظروف هذه الدراسة.

بالنسبة للإجابة على سؤال الدراسة الثالث والمتعلق بأثر كل من حجم العينة، وطول الاختبار، وشكل الاختبار وتفاعلاتها على دقة تقدير دالة معلومات الاختبار، أشارت النتائج الى أن حجم العينة الأصغر (250 مفحوص) كان متوسط الخطأ المعياري في التقدير هو الأقل وبالتالي دقة تقدير أعلى، والاختبار الأطول كانت دقة التقدير عنده أعلى وذلك لتدني قيم متوسط الخطأ، وفقرات الصواب والخطأ حققت متوسطات خطأ أقل، أي أنه يمكن القول بأن تطبيق الاختبار الأطول (50 فقرة) الذي فقراته على شكل صواب وخطأ على عينة مكوَنة من (250 مفحوص) يحقق دقة أعلى في تقدير دالة معلومات الاختبار مقارنة مع ظروف وحدود هذه الدراسة.

التوصيات:

في ضوء النتائج التي تم التوصل إليها في هذه الدراسة يقترح الباحث مجموعة من التوصيات تتعلق بموضوع هذه الدراسة وهي:

1) استخدام عيّنات كبيرة من المفحوصين وعدد ملائم من الفقرات – يفضل 50 فقرة – ومن نوع الاختيار من متعدد – عند مقارنتها باختبارات من نوع صواب وخطأ – للاختبارات ثنائية الإجابة عند إجراء عملية تقدير لمعالم الفقرة والقدرة ذلك لتحقيق أكبر قدر من الدقة في التقدير.

2) إجراء المزيد من الدراسات تتعلق بموضوع دقة تقدير المعالم ودالة معلومات الاختبار باستخدام أنواع وأشكال أخرى من الفقرات مثل الفقرات ذات الاجابة القصيرة، وفقرات المطابقة أو المزاوجة، واستخدام طرق تقدير أخرى، واستخدام مؤشرات أخرى لدقة التقدير مثل التحَيز والجذر التربيعي لمتوسط مربع الخطأ، وتطبيق ذلك على مواد دراسية أخرى غير الرياضيات مثل العلوم واللغة العربية وغيرها.

3) الاستفادة من الاختبارات التحصيلية التي تم إعدادها في هذه الدراسة من قبل معلمي المبحث، أو من قبل باحثين آخرين.

4) تناول بعض المتغيرات الأخرى، ودراسة تأثيرها في دقة تقدير معالم الفقرة والقدرة ودالة معلومات الاختبار مثل متغير جنس المفحوصين (ذكر، أنثى)، وعدد بدائل الإجابة، وطرق التقدير المختلفة، والنموذج اللوجستي المستخدم.

المراجـــــع

المراجع العربية

التقي، أحمد محمد (2009). **النظرية الحديثة في القياس،** الطبعة الأولى. عمان: دارالمسيرة.

الثوابية، أحمد (2010). أثر حجم العينة على تقدير صعوبة الفقرة والخطأ المعياري في تقديرها باستخدام نظرية الاستجابة للفقرة. **مجلة جامعة دمشق،** 26(1)، 525-556.

الحجيلي، خالد بن هديبان (2012). **أثر طول الاختبار وحجم العينة في دقة تقدير معلمة صعوبة الفقرة والقدرة ومعادلة الاختبارات بوجود الأداء التفاضلي للفقرة** .أطروحة دكتوراه غير منشورة، جامعة اليرموك، اربد، الأردن.

الدرابيع، ماهر (2001). فعالية النموذج اللوغاريتمي ذي المعلمة الواحدة "نموذج راش" في تقدير قدرة الفرد ومعامل صعوبة الفقرة بإختلاف حجم العينة وطول الإختبار. **مجلة دراسات، العلوم الإنسانية، 28**(1)، 197- 208.

الزبون، حابس (2013). أثر حجم العينة على تقدير دالة المعلومات للاختبار والخطأ المعياري في تقديرها باستخدام النظرية الحديثة في القياس. **مجلة جامعة النجاح للأبحاث،** العلوم الانسانية 27(6)، 1324-1344.

زيتون، عايش محمود (1996). **اساسيات الإحصاء الوصفي.** عمان: دار عمار.

الشريفين، نضال (2012). أثر طريقة تقدير معالم الفقرة وقدرات الأفراد على قيم معالم الفقرة، والخصائص السيكومترية للاختبار في ضوء تغير حجم العينة. **المجلة التربوية، 104**(1)، 238-177

عبابنة، عماد (2004). **أثر حجم العينة وطريقة انتقائها وعدد الفقرات وطريقة انتقائها على دقة تقدير معالم الفقرة، والقدرة لاختبار قدرة عقلية باستخدام نظرية الاستجابة للفقرة.** اطروحة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، عمّان، الأردن.

علام، صلاح الدين محمود (2005). **نماذج الاستجابة للمفردة الاختبارية أحادية البعد متعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي**، ط 1. القاهرة: دار الفكر العربي.

علام، صلاح الدين محمود (2001). **الاختبارات التشخيصية مرجعية المحك**، مصر: جامعة الأزهر، دار الفكر العربي.

عودة، احمد سليمان (2004). **القياس والتقويم في العملية التدريسية (ط1)**. إربد: دار الامل للنشر .

العنزي، محمد طالب (2010). **أثر شكل الفقرة على معالم الفقرة وثبات الاختبار وفقاً لنظرية الاستجابة للفقرة**، رسالة ماجستير غير منشورة، جامعة مؤتة، الكرك، الاردن.

المصري، أحمد محمود (2009). **أثر شكل فقرات الاختبار في دقة تقدير قدرات المفحوصين ومعالم الفقرات وفق نظرية استجابة الفقرة.** أطروحة دكتوراه غير منشورة، جامعة اليرموك، اربد، الأردن.

النصراوين، معين سلمان (2015). **مقارنة أربعة نماذج لمعالجة التخمين في فقرات الاختيار من متعدد في اطار النموذج اللوجستي ثلاثي المعلمة وأثرها على دقة تقدير معلمة القدرة ومعالم الفقرات**، أطروحة دكتوراه غير منشورة، الجامعة الأردنية، عمان، الاردن.

المراجع الأجنبية

Baker, F. (1985). **The Basics of Item Response Theory**. Portsmouth, NH: Heinman.

Baker, F. (1987). Methodology review: Item parameter estimation under the one, tow and three-parameter logistic models. **Applied Psychological Measurement, 11**(2), 111-141.

Barnes, L.,&Wise (1991). The utility of a modified one - parameter IRT model with small samples. **Journal of Applied Measurement in Education**, 4 (2), 143 – 153.

Bock, R. &Mislevy,R. (1982). Adaptive EAP estimation of ability in amicro computer environment. **Applied psychological Measurement,**6, 431-444.

Birnbaum, A. (1986). Some latent trait models and their use in inferring an examminees ability. In F. M . Lord & M. R. Novick (Eds), **Statistical theories of mental test scores.** Reading, MA: Addison-Wesley.

Casella, A., & Berger, R. (1986). **Statistical inference.** California: wadworth and Brooks / cole advanced books & software Pacific grove.

Crocker, L. & Algina, J. (1986). **Introduction to classical and modern test theory**. USA: Florida**,** University of Florida.

Dawber, Teresa. Rogers, W.Todd and Carbonaro, Michael.(2009). Robustness of Lord's Formulas for Item Difficulty and Discrimination Conversions Between Classical and Item Theory Models.**The Alberta Journal of Educational research**, 4(55), 512-533.

Dennick,Reg. & Tavacol, Mohsen. (2011). Making sense of Coronback's Alpha. **International Journal of Medical Education, 2,** 53 – 55.

Embretson, S., & Steven, Reise, (2000). **Item Response Theory for Psychologists.** New Jersy: Lawrence Erlbaum.

Glass, G. (2005).**The impact of item parameter estimation of computerized adaptive testing with item cloning**. Law School Admission Conceit Computerized Testing Report 02-06 November.

Gronlund, N., & Linn, R. (1990). **Measurement and evaluation in teaching.** New York: Macmillan Publishing Company.

Harlow, L. Lisa. (2005). **The Essence of Multivariate Thinking Basic Themes and Methods, Lawrence Erlbaum Associates,** New Jersey.

Hambelton, R., & Jones, R. (1993). Comparison of classical test theory and item response theory and their applications to test development, **Journal of Educational Measurement. Issues and Practice,** 38-47

Hambelton, R., & Swaminathan, H. (1985**). Item response theory principles and application**. Boston: Kluwer–Nigh of Publishing.

Hambelton,R. & Rogers (1988). Solving criterion referenced measurement problems with item response models. Paper Presented at **the Annual Meeting of the American Educational Research Association**, New Orleans.

Hambelton, R., Swaminathan, H., & Rogers, H. (1991). **Fundamentals of item response theory.** Newbury, New York :Sage.

Hambleton, R. , Cook L. (1983 **) .** Robutness of item response models and effects of test length and sample size in the precision of ability estimates . New York . In D.J . Weiss(Ed) **New horizons in testing**.

Hattie,J . (1985) **.** Methodology Review : Assessing Unidimensionality of Tests and Items . **Applied psychological Measurment**,9,139-164 .

Hopkins,C . and Antes, R. (1978) **. Classroom Measurement and Evaluation Itasca,** F.E . peacock publishers, Inc.

Lim, T.Hoon, B. Hardwell & Michael (1997**).** Effects of Number of Items and Examinees on Parameter Estimation in Item Response Theory, AResearch Synthesis. Paper Presented **at the Annual Meeting of the American Educational Research Association** .Chiocago,IL,March 24-28.

Lord, F.(1986) **.** Maximum Likelihood and Bayesion parameter estimation in item response theory**. Journal of Educational Measurement**, 23, 157-162 .

Lord, F.& Novick, M .(1968 **). Statistical Theories of Mental Test Scores .**N.Y : Addison – Wesley .

Lord, F.(1980**). Applications of Items Response Theory to Practical TestingProblems**. Hillsdale, N.J,Lawrence Erlbaum.

Mislevy, R.& Bock,R.(1997). Bilog 3. **Item analysis and test scoring with binary Logistic models.** Chicago, I L: SSI scientific software international.

Mislevy, R., & Bock, R. (2003)**. BILOG (version3) [computer software] .** Lincoln wood, IL: Scientific Software International .

Pelton, W. (2002). **The accuracy of unidimensional Measurement Models in the presence of deviations from the underlying assumptions.**

Unpublished doctoral dissertation, Brigham young university, Provo, Utah, USA.

Qing Ping He, Chris Wheadon (2013). The Effect of Sample Size on Item Parameter Estimation for Partial Credit Model. **Journal of Educational Measurement**, 27(2). 165-174.

Thissen, D. & wainer, H .(1982) **.** Some standard errors in item response theory **. Psychometrika**, 47,(4) 154-161.

Yen,w.(1993) . Scaling Performance Assessments : Strategies for Managing Local Item Dependence. **Journal of Educational Measurement**,30,187-213.

ملاحق الدراسة

ملحق (1)

أسماء المدارس والجنس وعدد الصفوف وعدد الطلبة في كل صف في عينة الدراسة

الرقم	اسم المدرسة	الجنس	الشعبة	العدد	المجموع
1	منذر المصري	ذكور	أ	30	**146**
			ب	29	
			جـ	25	
			هـ	21	
			و	23	
			ز	18	
2	عائشة أم المؤمنين	اناث	أ	27	**157**
			ب	33	
			جـ	30	
			د	35	
			هـ	32	
3	رقية بنت الرسول	اناث	أ	36	**132**
			ب	33	
			جـ	37	
			د	26	
4	عاكف الفايز	ذكور	أ	38	**110**
			ب	37	
			جـ	35	
5	الأشرفية الثانوية	اناث	أ	25	**78**
			ب	26	
			جـ	27	
6	حي أم تينة	اناث	أ	24	**94**
			ب	24	
			جـ	22	
			د	24	
7	زهران الثانوية	ذكور	أ	27	**93**
			ب	37	
			جـ	29	
8	الشميساني الغربي	اناث	أ	32	**64**
			ب	32	
9	الأمير راشد بن الحسن	ذكور	أ	27	**81**
			ب	30	
			جـ	24	
10	الملكة نور الحسين	اناث	أ	27	**60**
			جـ	33	

ملحق (2)
تحليل محتوى مادة الرياضيات (الفصل الأول) للصف الثامن الأساسي

عنوان الوحدة	المفاهيم والمصطلحات	الحقائق والأفكار	الرموز	المهارات	القيم والإتجاهات
الوحدة الأولى: الأعداد الحقيقية	العدد الصحيح، العدد النسبي، العدد الحقيقي، الجذوربأنواعها،الخاصيتين : التبديلية، التجميعية،العنصرالمحايد، النظيرالجمعي والنظير الضربي، انطاق المقام، الفترات بأنواعها، المتباينة والمعادلة، المتباينة الخطية، الأسس، الأسس النسبية، الصورة العلمية للعدد.	* كل عدد يكتب على صورة $\frac{أ}{ب}$ فهو عدد نسبي، وغير ذلك فهو عدد غير نسبي. * كل عددحقيقي هوعدد نسبي. * مجموعات الأعداد: الطبيعية،الصحيحة، النسبية وغير النسبية هي مجموعات متداخلة ضمن نظام الأعداد الحقيقية. * تبسيط التعابير الجذرية. *خواص نظام الأعداد الحقيقية. *القوانين المتعلقة بالأسس الصحيحة. *قواعد الأسس النسبية للأعداد الصحيحة. *الأعداد الكبيرة جداً والأعداد الصغيرة جداً، والصيغة العلمية. *المسائل العملية على مجموعة الأعداد	*القوى: $س^2$، $س^3$. * الجذر التربيعي والجذر التكعيبي. * باي (π). * رموزالفترات بأنواعها: المغلقة، المفتوحة، نصف المغلقة (أونصف المفتوحة). * رمزيَ المالانهاية (∞، - ∞). * جذور مثل $\sqrt[ن]{س}$.	* يصنف الأعداد الحقيقية الى مجموعات: أعداد طبيعية، وصحيحة، ونسبية، وغير نسبية. * يبسط التعابير الجذرية. * يستخدم القوانين المتعلقة بالأسس الصحيحة. * يطبق القوانين المتعلقة بالأسس الصحيحة، والقوى النسبية للأعداد الصحيحة. * يتقن التعامل مع الأعداد الكبيرة جداً والصغيرة جداً. * يستخدم الرموز في الآلة الحاسبة بشكل صحيح. * القدرة على حل مسائل حياتية تتعلق بمجموعة الأعداد الحقيقية.	*يميّز بدقة بين مجموعات الأعداد. * يتحدث بلغة رياضية سليمة. * يكتب الجذور والأسس بشكل رياضي صحيح. * يوضح بلغته الخاصة القوانين المتعلقة بالأسس الصحيحة والنسبية. * يطوّر فهماً للأعداد الصغيرة والكبيرة باستخدام القوى. * توسيع مداركه وفهمه للأعداد بمجموعاتها المختلفة.

		الحقيقية.			
الوحدة الثانية: المقادير الجبرية	الحد الجبري، المقدار الجبري، أولويات العمليات الحسابية، الحدود الجبرية المتشابهة، جمع الحدود الجبرية وطرحها، ضرب المقادير الجبرية، الضرب الأفقي والضرب العمودي، العامل المشترك الأكبر (ع.م.أ)، التحليل الى العوامل بإخراج العامل المشترك.	*الحد الجبري: يتكون من حاصل ضرب عدد ثابت بمتغيّر أو أكثر. * الحدود الجبرية المتشابهة التي لها نفس القسم الرمزي وإن اختلفت المعاملات. * جمع وطرح معاملات الحدود الجبرية المتشابهة. * ضرب الحدود الجبرية بضرب المعاملات وضرب القسم الرمزي للحدود الجبرية. * توزيع الضرب على الجمع. * ضرب المقادير الجبرية. * العامل المشترك الأكبر لحدين أو أكثر. * التحليل بالتجميع وإخراج العامل المشترك.	العامل المشترك الأكبر(ع.م.أ)، المتغيرات (س،ص،ع،...)، $(أ+ب)^2 = أ^2 + 2أب + ب^2$. أ×(ب+ج)=أب+أج.	*يجري عمليات الجمع، الطرح، الضرب) على المقادير الجبرية بطريقة صحيحة ومتقنة. * يخرج العامل المشترك الأكبر لحدود جبرية. * يحلل المقدار الجبري بالتجميع واخراج العامل المشترك الأكبر بدون أخطاء.	* يميّزبين الحدود الجبرية والمقادير الجبرية. * يتقن اجراء العمليات الحسابية الثلاث (جمع، طرح، ضرب) على الحدود الجبرية. * يتقن ايجاد العامل المشترك الأكبر. * يحلل المقادير الجبرية بإخراج العامل المشترك. * يوظف ما تعلمه في هذه الوحدة بالنسبة للحدود والمقادير الجبرية في حل مشكلات ومسائل حياتية.
	الأنماط، العلاقة، الإقتران، المخطط السهمي، المجال، المدى، بيان الإقتران، اختبار الخط الرأسي، الإقتران الخطي، الإقتران الثابت، الاقتران المتزايد، الإقتران		*رمز المجموعة { }. *الأزواج المرتبة. *رمز الإقتران ق(س).	*يمثل الإقتران بمخطط سهمي وبياني. * يتقن إيجاد مجال الإقتران ومداه. *يميّز الإقتران من العلاقة مستخدماً اختبار	

الوحدة الثالثة: الأنماط والإقترانات	المتناقص، ميل منحنى الإقتران، المقطع الصادي للإقتران الخطي، معامل س. *اكتشاف قاعدة النمط. * تعريف الإقتران. *الإقتران الثابت: ق(س)= ب. * قاعدة الإقتران الخطي: ق(س)= أس + ب، ميل الإقتران=معامل س=أ *ق(س) متزايد إذا كانت أ> 0 *ق(س) متناقص إذاكانت أ<0 *مقطع الخط المستقيم من محور الصادات= الحد الثابت(ب).	*المجال والمدى.	الخط الرأسي. *يميّزبين الإقتران الخطي المتزايد والإقتران الخطي المتناقص. *يعمم الطالب الأنماط في مواقف حل المشكلات. * يميّز الطالب الإقتران من خلال العلاقة. * يجد قاعدة الإقتران الخطي. * يطوّر القدرة والمهارة في تمييز الإقتران الخطي بيانياً. * يمتلك القدرة على التعرف الى خواص الإقتران الخطي(التزايد والتناقص، المقطع الصادي، الميل). * يحل مسائل حياتية.

ملحق (3)

النتاجات التعليمية (اهداف المحتوى) للفصل الأول لمادة الرياضيات للصف الثامن الأساسي

- **الوحدة الأولى: الأعداد الحقيقية**

الهدف(1) : يصنف الأعداد الحقيقية كمجموعات: أعداد طبيعية وصحيحة ونسبية وغير نسبية.

الهدف(2): يبسّط التعابير الجذرية التي تتضمن أعداداً حقيقية.

الهدف(3): يفهم خواص نظام الاعداد الحقيقية.

الهدف (4): يذكر خصائص عمليتي الجمع والضرب على الأعداد الحقيقية.

الهدف (5) : يتعرف على الفترات في مجموعة الاعداد الحقيقية.

الهدف (6) : يحل متباينة خطية بمتغير واحد.

الهدف(7): يستخدم القوانين المتعلقة بالاسس الصحيحة.

الهدف (8) : يعبر عن الاعداد الكبيرة والاعداد الصغيرة بالصورة العلمية.

الهدف(9): يعبرعن بعض الأعداد الكبيرة وكسورها ببعض الألفاظ والرموزالإغريقية.

الهدف(10): يستخدم القوانين المتعلقة بالأسس النسبية.

الهدف(11): يكتب بعض الأعداد الحقيقية في أبسط صورة.

الهدف(12): يحل مسائل عملية على الأعداد الحقيقية.

- **الوحدة الثانية: المقادير الجبرية**

الهدف(13) : يتعرف الحدود والمقادير الجبرية.

الهدف(14) : يجد القيمة العددية للمقدار الجبري.

الهدف (15) : يميّز الحدود الجبرية المتشابهة.

الهدف(16): يجد ناتج جمع الحدود الجبرية وطرحها.

الهدف(17) : يجد ناتج ضرب حد جبري في حد جبري اخر.

الهدف(18) : يجد ناتج ضرب حد جبري في مقدار جبري اخر.

الهدف(19) : يجد ناتج ضرب مقدار جبري في مقدار جبري اخر.

الهدف(20): يجد العامل المشترك الأكبرللحدود الجبرية.

الهدف(21): يحلل المقدار الجبري بإخراج العامل المشترك الاكبر.

الهدف (22): يحل مسائل عملية تتعلق بالمقادير الجبرية.

- **الوحدة الثالثة : الانماط والاقترانات**

الهدف(23) : يعمم الانماط في مواقف حل المشكلات.

الهدف(24) : يميّزالاقتران من العلاقة.

الهدف(25) : يستخدم اختبار الخط الرأسي لتمييز الاقتران.

الهدف(26): يجد قاعدة الاقتران الخطي.

الهدف(27): يتعرف الى الاقتران الخطي.

الهدف(28): يجد صورة عدد حقيقي في اقتران خطي.

الهدف(29): يستوعب خواص الاقتران الخطي.

الهدف (30): يحل مسائل حياتية على المقادير الجبرية.

ملحق (4)

الاختبار في صورته الأولية (المكوَن من 96 فقرة) والذي عُرض على المحكمين.

ملاحظة(1): كل هدف يليه مجموعة من الفقرات التي تقيسه ليخضع لتقييم المحكمين.
ملاحظة(2): تم وضع الإشارة(*) فوق البديل الصحيح لكل فقرة.

اختبار تحصيلي في الرياضيات للصف الثامن الاساسي (الفصل الاول) من نوع الاختيار من متعدد

- **الوحدة الأولى: الأعداد الحقيقية**

الهدف(1) : يصنف الأعداد الحقيقية كمجموعات: أعداد طبيعية وصحيحة ونسبية وغير نسبية[فهم واستيعاب]

1) **واحد من الاعداد الآتية يعد عدد غير نسبي:**

(أ) 8   (ب) -8   (ج*) $\sqrt{8}$ (د) $0,\overline{8}$

2) **واحدة من العبارات الاتية صحيحة:**

(أ) كل عدد حقيقي هو عدد نسبي   (ب*) كل عدد نسبي هو عدد حقيقي

(ج) كل عدد حقيقي هو عدد غير نسبي   (د) كل عدد نسبي هو عدد صحيح

3) العدد0, 151151115000 ينتمي الى مجموعة الاعداد:

(أ) الصحيحة   (ب) الطبيعية   (ج) النسبية   (د*) الحقيقية

4) واحد من الاعداد الاتية <u>لا يمكن</u> كتابته على الصورة $\frac{أ}{ب}$:

(أ) 0,65   (ب*) $\sqrt{23}$ (ج) 23   (د) $0,\overline{23}$ (23 دوري)

5) واحد من الاعداد الاتية <u>يمكن</u> كتابته على صورة $\frac{أ}{ب}$:

(أ*) $0,\overline{3}$ (ب) π   (ج) $\sqrt{2}$ (د) 4,6767767770000

الهدف(2): يبسّط التعابير الجذرية التي تتضمن أعداداً حقيقية [فهم واستيعاب]

6) قيمة المقدار $\sqrt[3]{-64}$ تساوي:

(أ) -8   (ب) 8   (ج*) -4   (د) 4

7) قيمة المقدار $\sqrt{0.16}$ =

أ) 4   ب*) 0,4   ج) 0,8   د) 8

8) عند كتابة العدد الكسري $2\frac{7}{9}$ على صورة عدد مرفوع للأس 2 يصبح:

أ*) $(\frac{5}{3})^2$ ب) $(\frac{3}{5})^2$ ج) $(\frac{7}{9})^2$ د) $(2\frac{7}{9})^2$

9) قيمة المقدار $\sqrt[3]{\frac{27}{64}}$ =

أ) $\frac{9}{8}$ ب) $-\frac{9}{8}$ ج) $\frac{27}{64}$ د*) $\frac{3}{4}$

10) $\sqrt[3]{س^3}$ =

أ) $س^3$ ب*) س   ج) $س^2$ د) 3

11) $\sqrt[3]{15\frac{5}{8}}$ =

أ) $\frac{5}{8}$ ب) $\frac{125}{8}$ ج*) $\frac{5}{2}$ د) $\frac{2}{5}$

الهدف(3): يفهم خواص نظام الاعداد الحقيقية [فهم واستيعاب]

12) $\sqrt{س \times ص}$

أ) س×ص   ب) $\sqrt{س} \times ص$ ج) $س \times \sqrt{ص}$ د*) $\sqrt{س} \times \sqrt{ص}$

13) قيمة المقدار $\sqrt[3]{125 \times 64}$ =

أ*) 20   ب) 40   ج) 0,20   د) 0,40

14) $\sqrt[3]{\frac{-8}{27}}$ =

أ) $\frac{2}{3}$ ب) $\frac{4}{9}$ ج*) $\frac{-2}{3}$ د) $\frac{-4}{9}$

15) واحدة من العبارات الاتية خاطئة:

(أ) $\sqrt{\frac{25}{9}}=\frac{5}{3}$ (ب*) $\sqrt{9+25}=8$ (ج) $\sqrt{9\times25}=15$ (د) $\sqrt{9-25}=4$

16) $\sqrt{3}+5\sqrt{3}=$

(أ*) $6\sqrt{3}$ (ب) $5\sqrt{9}$ (ج) $5\sqrt{6}$ (د) $6\sqrt{6}$

17) $\sqrt{18}+\sqrt{2}=$

(أ) $\sqrt{20}$ (ب) 20 (ج*) $4\sqrt{2}$ (د) $2\sqrt{18}$

الهدف (4): يذكر خصائص عمليتي الجمع والضرب على الأعداد الحقيقية. [معرفة وتذكر]

18) (س+ص)+ع = س+(ص+ع) تسمى هذه الخاصية على الاعداد الحقيقية بالخاصية:

(أ) التبديلية (ب*) التجميعية (ج) العنصر المحايد لعملية الجمع (د) النظير الجمعي

19) النظير الجمعي للعدد ($-\sqrt{17}$) هو:

(أ) $\frac{1}{\sqrt{17}}$ (ب) $\frac{1}{17}$ (ج) 17 (د*) $\sqrt{17}$

20) العنصر المحايد لعملية الجمع على مجموعة الاعداد الحقيقية هو:

(أ*) صفر (ب) 1 (ج) -1 (د) 5

21) النظير الضربي للعدد $\frac{3}{4}$ هو:

(أ) $\frac{-3}{4}$ (ب) 1 (ج*) $\frac{4}{3}$ (د) صفر

22) $\sqrt{3}\,(\sqrt{3}+\sqrt{12})=$

(أ) 3 (ب*) 9 (ج) $3\sqrt{12}$ (د) $3\sqrt{15}$

23) عند انطاق المقام للعدد $\frac{1}{\sqrt{10}+3}$ بحيث لا يظهر الجذر في المقام يصبح:

(أ) $\sqrt{10}+3$ (ب) $\frac{1}{3}$ (ج*) $\sqrt{10}-3$ (د) $\frac{1}{\sqrt{10}}$

الهدف (5) : يتعرف على الفترات في مجموعة الاعداد الحقيقية [معرفة وتذكر]

24) واحدة من الفترات الاتية تعد فترة مغلقة:

(أ) (أ ، ب) (ب) [أ ، ب) (ج) (أ ، ب] (د*) [أ ، ب]

25) اذا كانت المجموعة : ف ={ س : س ϶ ح -4< س ≥ 3 } فإن التعبير عنها باستخدام رمز الفترة يصبح:

(أ) [-4 ، 3] (ب*) (-4 ، 3] (ج) [3 ، -4] (د) (-4 ،3)

26) عند ضرب طرفي المتباينة $\sqrt{3} < 2$ بالعدد -1 تصبح:

(أ) - $\sqrt{3}$ > 2 (ب) - $\sqrt{3}$ > -2 (ج*) - $\sqrt{3}$ < -2 (د) - $\sqrt{3}$ < 2

27) إحد من الفترات الاتية تسمى فترة غير محدودة:

(أ) (أ ، ب) (ب) [أ ، ب] (ج*) [أ ، ∞) (د) (أ ، ب]

28) طول الفترة [-2 ، 3] يساوي:

(أ) 1 (ب) -1 (ج) -5 (د*) 5

29) اذا كانت المجموعة : ف ={س : س ϶ح ، س ≤ 3 } فإن التعبير عنها باستخدام رمزالفترة تصبح:

(أ*) [3 ، ∞) (ب) (3 ، ∞) (ج) [3 ، - ∞) (د) (∞ ، 3]

الهدف (6) : يحل متباينة خطية بمتغير واحد [تطبيق]

30) مجموعة حل المتباينة : س- 7 > 5 هو:

(أ) (-2 ، ∞) (ب) [12 ، ∞] (ج*) (12 ، ∞) (د) (- ∞ ، 2)

31) مجموعة حل المتباينة $\frac{3}{4}$س ≥ $\frac{7}{8}$ هو:

(أ*) (-∞ ، $\frac{7}{6}$] (ب) (-∞ ، $\frac{7}{6}$) (ج) ($\frac{3}{4}$ ، $\frac{7}{8}$] (د) (-∞ ، ∞)

32) مجموعة حل المتباينة 3س + 7 > 9- 4س هو:

(أ) (-∞ ، $\frac{1}{3}$) (ب) (- $\frac{1}{3}$ ، $\frac{1}{3}$) (ج*) ($\frac{1}{3}$ ، ∞) (د) [$\frac{1}{3}$ ، ∞)

33) قيمة س التي تحقق المعادلة: $\frac{1}{\sqrt{3}}$ س = $\sqrt{15}$ هي:

(أ) 45 (ب*) 3$\sqrt{5}$ ج) $\sqrt{3}$ د) 3$\sqrt{3}$

الهدف(7): يستخدم القوانين المتعلقة بالاسس الصحيحة [تطبيق]

34) $((\sqrt{5})^3)^2$ =

(أ*) 125 (ب) 225 (ج) 25 (د) 5

35) $(\sqrt{12})^{-4} =$

(أ) 144 (ب) $\sqrt{12}$ (ج*) $\frac{1}{144}$ (د) - $\frac{1}{144}$

36) $\frac{(\sqrt{5}-1)^{15}}{(\sqrt{5}-1)^{13}} =$

(أ) $3 - \sqrt{5}$ (ب*) $6 - 2\sqrt{5}$ (ج) $\sqrt{5} - 1$ (د) 1

الهدف (8) : يعبر عن الاعداد الكبيرة والاعداد الصغيرة بالصورة العلمية [فهم واستيعاب]

37) يكتب العدد (4083000000) على الصورة العلمية على النحو:

(أ) $4,083 \times 10^{-9}$ (ب) $3, 408\times10^{9}$ (ج) $40,83 \times 10^{9}$ (د*) $4,083\times10^{9}$

38) يكتب العدد (0٫0000423) على الصورة العلمية:

(أ) $4,23 \times 10^{5}$ (ب*) $4,23\times10^{-5}$ (ج) $42,3\times10^{5}$ (د) $42,3\times10^{-5}$

الهدف(9): يعبرعن بعض الأعداد الكبيرة وكسورها ببعض الألفاظ والرموزالإغريقية. [معرفة وتذكر]

39) (1) جيجا (giga) يساوي:

(أ) مليون (10^{6}) (ب) مليون مليون(10^{12}) (ج*)الف مليون(10^{9})
(د)مئة مليون(10^{8})

40) (1) ميكرو(micro) يساوي:

(أ) جزء من عشرة(10^{-1}) (ب*) جزء من مليون(10^{-6})
(ج) جزء من الف مليون(10^{-9}) (د) جزء من الف (10^{-3})

الهدف(10): يستخدم القوانين المتعلقة بالأسس النسبية [تطبيق]

41) $(64)^{1/3} =$

(أ*)4 (ب) 8 (ج) -4 (د) -8

42) $(-128)^{1/7} =$

(أ) 2 (ب) 3 (ج*) -2 (د) -3

43) $(\sqrt{17})^{4/3} \times (\sqrt{17})^{5/4} =$

(أ) -17 (ب) $\sqrt{17}$ (ج) $\sqrt[3]{17}$ (د*) 17

الهدف(11): يكتب بعض الأعداد الحقيقية في أبسط صورة. [تطبيق]

44) عند كتابة المقدار $\sqrt[5]{\frac{-243}{32}}$ في أبسط صورة يصبح:

(أ) $\frac{3}{2}$ (ب) $\frac{2}{3}$ (ج*) $-\frac{3}{2}$ د) $-\frac{2}{3}$

45) عند اختصار المقدار $\frac{س^{4/3} \times ص^{3/2}}{س^{2/1} \times ص^{6/1}}$ الى أبسط صورة يصبح:

(أ) $س^{6/4} ص^{9/5}$ (ب) $س ص^{-3/1}$ ج) $س^{2/1} ص^{4/1}$ (د*) $س^{4/1} ص^{2/1}$

46) $(\sqrt{3} \times \sqrt[3]{9})^6 =$

(أ*) 2187 (ب) 729 (ج) 243 (د) 2180

الهدف(12): يحل مسائل عملية على الأعداد الحقيقية. [تطبيق]

47) اشـــترى تـــاجر عـــدد مـــن لُعـــب الأطفـــال (س) بمبلـــغ (437) دينـــار، ويبيـــع اللُعبـــة الواحـــدة بمبلـــغ(5) دنـــانير، فـــإن أقـــل عـــدد مـــن لُعـــب الأطفـــال التـــي يجـــب أن يبيعهـــا حتـــى يحقق ربح هو:

(أ) 87 (ب) 86 (ج*) 88 (د) 89

48) غرفـــة مربعـــة الشـــكل مســـاحتها = $40\frac{96}{100}$ م^2، فـــإن طـــول الســـجادة التـــي تغطـــي الغرفة تماماً:

(أ) 64م (ب) 4,6 م (ج*) 6,4 م (د) 64,0

49) مســـتطيل طولـــه $\sqrt{360}$ ســـم، وعرضـــه $\sqrt{40}$ ســـم، فـــإن مســـاحته بالســـنتيمتر المربـــع بأبسط صورة تساوي:

(أ*) 120 (ب) 20 (ج) 14400 (د) 12

❖ **الوحدة الثانية: المقادير الجبرية**

الهدف(13) : يتعرف الحدود والمقادير الجبرية [معرفة وتذكر]

50) القسم الرمزي للحد الجبري ($\sqrt{5}$ س^{3}ص ع) هو:

(أ) $\sqrt{5}$ (ب) س^3 (ج) س^{3}ص (د*) س^{3}ص ع

51) معامل الحد الجبري ($\frac{5}{7}$ س ص) هو:

(أ) $\frac{7}{5}$ (ب*) $\frac{5}{7}$ (ج) س ص (د) $\frac{5}{7}$ س

52) حديقة على شكل مستطيل طولها (س) متراً وعرضها (ص) متراً فإن المقدار الجبري الذي يمثل محيطها

(أ) س ص (ب) س+ص (ج) 2س +ص (د*) 2س + 2ص

الهدف(14) : يجد القيمة العددية للمقدار الجبري [تطبيق]

- **اذا كانت س =5، ص = -3، ع =6 ، فأجب عن السؤالين (46، 47)، الاتيين:**

53) القيمة العددية للمقدار الجبري (7س ص) يساوي :

(أ*) – 105 (ب) 105 (ج) 35 (د) -21

54) القيمة العددية للمقدار (2س ع + ص^2) يساوي:

أ) 51 (ب*)69 (ج) -30 (د) 60

الهدف (15) : يميّز الحدود الجبرية المتشابهة [معرفة وتذكر]

55) الحد الجبري الذي يشابه الحد الجبري (7س^{3}ع) هو:

(أ) 7س ع (ب) س ع^3 (ج*) 8س^{3}ع (د) 6س^{3}ع^3

الهدف(16): يجد ناتج جمع الحدود الجبرية وطرحها [فهم واستيعاب]

56) 2س + 5س =

(أ) 7س^2 (ب*) 7 (ج) 7 (د) س^2

57) (6س + 3ص) + (5ع – 7ل) =

(أ) 11س ع + 4ص (ب)6س+3ص+5ع (ج*)6س+3ص+5ع-7ل (د)9س ص – 2ص ل

58) (12ص+4س^2-3ل) – (7س^2+5ص) =

(أ) 13ص س^2 ل (ب)7ص- 3س^2+3ل (ج)5ص+9س-3ل (د*)7ص-3س^2-3ل

الهدف(17) : يجد ناتج ضرب حد جبري في حد جبري اخر[فهم واستيعاب]

59) 4ص^2 ع^3 ×3ع^2 ص^3 =

(أ*)12ص^5 ع^5 (ب)12ص^2 ع^5 (ج)12ص^{5}ع^3 (د)12ص^4 ع^6

الهدف(18) : يجد ناتج ضرب حد جبري في مقدار جبري اخر[فهم واستيعاب]

60) 8ص(ص+3ع) =

(أ)8ص^2+3ع (ب)8ص+24ص ع (ج*)8ص^2+24ص ع (د)8ص^2+8ص ع

61) -3×(4-5أ) =

(أ) -12-15أ (ب)-12+15أ^2 (ج)12+15أ (د*)-12+15أ

الهدف(19) : يجد ناتج ضرب مقدار جبري في مقدار جبري اخر [تطبيق]

62) (3س+2)(5س+3ص)=

(أ*)15س^2+9س ص+10س+6ص (ب)15س^2+6ص

(ج)15س^2+10س+6ص (د)15س+9س ص+10س

63) صندوق مكعب الشكل طول حرفه(س^2+5)، يعبر عن حجمه باستخدام الأسس كما يأتي:

(أ) (س^2+5)^2 (ب*) (س^2+5)^3 (ج) 2(س^2+5) (د) 3(س^2+5)

64) (س+ص)^2=

(أ) س^2+ص^2 (ب)س^2- 2س ص+ص (ج*) س^2+2س ص+ص^2

(د)س+2س ص+ص^2

الهدف(20): يجد العامل المشترك الأكبرللحدود الجبرية [فهم واستيعاب]

65) العامل المشترك الاكبر للحدين (3س^2 ، 6س)هو:

(أ) 2س (ب) 18س^3 (ج) 9س^3 (د*) 3س

66) ع.م.أ للحدود الجبرية : 3س ص^2 ، 15س^{2}ص ، 27س^{2}ص^3 هو:

(أ*) 3س ص (ب) 15س ص (ج) 3س^{2}ص (د) 3س ص^2

الهدف(21): يحلل المقدار الجبري بإخراج العامل المشترك الاكبر [تطبيق]

67) عند تحليل المقدار الجبري (8أ^{2}ب – 10أ ب^2) يصبح:

(أ) 2أب(4أ+5ب) (ب) 2أب(أ^2 –ب^2) (ج) 2أب(ب^2-أ^2) (د*) 2أب(4أ- 5ب)

68) عند تحليل المقدار الجبري (8س^{2}ص^3+18ص^{2}س^3) بإخراج العامل المشترك يصبح:

(أ*) 2س^{2}ص^2(4ص+9س) (ب)8س^{2}ص^2(ص+8س)

(ج)2س^{2}ص^2(ص^3+18س) (د)2س ص(4ص+9س)

69) عند تحليل المقدار 15س ص – 4ل م +6ل ص –10س م بالتجميع وإخراج العامل المشترك يصبح:

(أ) (3ص+2ل)(5س+2م) (ب*) (3ص –2م)(5س+2ل)

(ج) (2م –2ص)(5س+2ل) (د) (5س – 2ل)(3ص+2م)

الهدف (22): يحل مسائل عملية تتعلق بالمقادير الجبرية [تطبيق]

70) خزان ماء على شكل متوازي مستطيلات، طوله (س+2)متراً، وعرضه يقل عن طوله بمقدار (2) متر، وارتفاعه ثلاثة أمثال عرضه فإن سعة الخزان (حجمه) بدلالة س هي:

(أ*) 3س^2(س+2) (ب) 3س(س+2) (ج) (س+2)(س –2) (د) 3(س+2)

71) قطعة أرض مستطيلة الشكل عرضها يساوي 2س متر، وطولها يساوي ثلاثة أمثال عرضها فإن مساحتها بدلالة س هي:

(أ) 12س (ب) 16س (ج*)12س^2 (د) 6س^2

❖ **الوحدة الثالثة : الانماط والاقترانات**

الهدف(23) : يعمم الانماط في مواقف حل المشكلات [تطبيق]

- **يحتوي خزان على 625م^3 من الماء، وينقص حجم الماء كل يوم بمقدار 25م^3 عن اليوم الذي قبله أجب عن السؤالين الآتيين.**

72) ان كمية الماء في الخزان المتبقية في اليوم الثاني تساوي بالمتر المكعب:

(أ) 600 (ب*) 575 (ج) 550 (د) 625

73) اذا كان عدد الايام = س فإن العلاقة التي تحدد كمية الماء في الخزان بدلالة عدد الايام هي:

(أ) 625+25س (ب) 625×25س (ج) 625÷25س (د*) 625 – 25س

الهدف(24) : يميّزالاقتران من العلاقة [معرفة وتذكر]

74) واحدة من العبارات الاتية صحيحة:

(أ) كل علاقة اقتران (ب) بعض الاقترانات علاقات (ج*) كل اقتران علاقة (د) العلاقة ليس لها مجال

75) القيم المدخلة وتمثل المساقط الاولى في الازواج المرتبة التي تمثل الاقتران تسمى:

(أ*) المجال (ب) المدى (ج) قاعدة الاقتران (د) بيان الاقتران

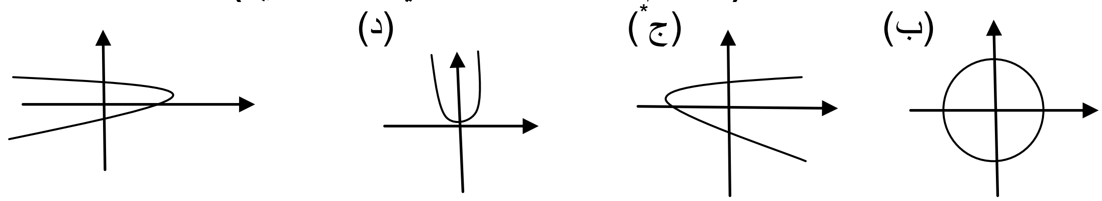
76) القيم الناتجة وتمثل المساقط الثانية في الازواج المرتبة التي تمثل الاقتران تسمى:

(أ) المجال (ب*) المدى (ج) قاعدة الاقتران (د) بيان الاقتران

الهدف(25) : يستخدم اختبار الخط الرأسي لتمييز الاقتران [تطبيق]

77) أي من العلاقات الاتية تمثل اقتران (استخدم اختبار الخط الرأسي أو العمودي)

(أ) (ب) (ج*) (د)



78) واحدة من العلاقات الاتية تمثل اقتران:

(أ*) {(-1،5)،(0،2)(3،1)} (ب) {(2،1)،(5،1)،(4،2)}

(ج) {(3،3)،(1،2)،(0،3)} (د) {(1،2)،(5،3)،(4،2)}

الهدف(26): يجد قاعدة الاقتران الخطي [تطبيق]

79) قاعدة الاقتران ق(س) الذي يمثله الجدول الاتي هي:

س	5	7	9	11
ص	9	13	17	21

(أ) ق(س)=2س (ب) ق(س)= س^2 (ج) ق(س)=2س+1 (د*) ق(س)= 2س -1

الهدف(27): يتعرف الى الاقتران الخطي [معرفة وتذكر]

80) أي مما يأتي يمثل اقتران خطي:

(أ) ق(س)= $\sqrt{س}$ +3 (ب*) ل(س)=2س-3 (ج)هـ(س)=4س$^{-2}$-3 (د) م(س)=5س^{2}+3

81) الصورة العامة للإقتران الخطي هي:

(أ*) ق(س)=أس+ب (ب)ل(س)= أس×ب س (ج)هـ(س)=اس^{2}+ب (د) ك(س)= $\frac{1}{س}$ +ب

82) واحد من الاقترانات الاتية يعد اقتراناً ثابتاً:

(أ) ق(س)= 5س+2 (ب) ل(س)= 8-س (ج*) م(س)= $\sqrt{3}$ (د) هـ(س)= 4+2س

83) عند تمثيل الاقتران ق(س)= 5س – 3 فإن الشكل الناتج يمثل :

(أ*) خط مستقيم (ب) نقطة (ج) منحنى (د) دائرة

الهدف(28): يجد صورة عدد حقيقي في اقتران خطي. [تطبيق]

84) اذا كان ق(س)=3س-2 فإن ق($\frac{1}{3}$)=

(أ) 7 (ب) 1 (ج) -2 (د*) -1

الهدف(29): يستوعب خواص الاقتران الخطي. [فهم واستيعاب]

85) ميل منحنى الاقتران ق(س) = 7-4س هو:

(أ) 4 (ب*) -4 (ج) -7 (د) 3

86) منحنى الاقتران ق(س)= 6س+9 يقطع محور الصادات عند النقطة:

(أ) (6،0) (ب) (15،0) (ج*) (9،0) (د) (9،6)

87) منحنى الاقتران ق(س)= 4س-6 يكون:

(أ*) متزايد (ب) متناقص (ج) ثابت (د) مقعر

88) منحنى الاقتران هـ(س)= 6-3س يكون:

(أ) ثابت (ب) متزايد (ج) مقعر (د*) متناقص

89) مدى الاقتران ق(س)={(-3،-1)،(-2،1)،(2،0)،(0،1)} هو:

(أ) {-3،-2،0،1} (ب*) {-1،2،0} ج) {-3،-2،0}

د) {-1،1،2}

90) ميل الاقتران ق(س)=12 يساوي:

(أ) 12 (ب) 1 (ج*) صفر (د) ∞

91) إذا كان ق(س)= 5س ـ 1، وكان ق(س)=9، فإن قيمة س =

(أ) 44 　　(ب*) 2 　　(ج) $\frac{8}{5}$ 　　(د) $\frac{9}{5}$

92) إذا كان ق(س) اقتران خطي ميله = 3، ومقطع منحناه من محور الصادات = ـ3، فإن قاعدة هذا الإقتران:

(أ*) ق(س) = 3س -3 　　ب) ق(س)= - 3س+3 　　(ج)ق(س)= 3س+3 　　(د)ق(س)= ـ 3س ـ 3

الهدف (30): يحل مسائل حياتية على المقادير الجبرية. [تطبيق]

93) أراد عامل اكمال بناء سور من الحجر، فوجد انه تم بناء (95) حجر قبل أن يبدأ بالعمل، فإذا قام العامل ببناء (23) حجر يومياً حتى أكمل بناء السورخلال (س) من الأيام فإن عدد الحجارة التي يتكون منها السور هي:

(أ) 95-23س 　　(ب) 23س - 95 　　(ج*) 95+23س 　　(د) 95×23س

94) إذا كان ثمن تذكرة دخول مدينة ألعاب ترفيهية دينار للفرد الواحد، وأجرة الاشتراك في أي لعبة من الألعاب (50 قرشاً) فإن المبلغ الذي يدفعه أحمد إذا اشترك في (6) ألعاب هو:

(أ) 3دنانير 　　(ب) 6دنانير 　　(ج) 7 دنانير 　　(د*) 4 دنانير

95) يراد احاطة حديقة مربعة الشكل، طول ضلعها(س) متراً، بسياج، تكلفة المترالواحد منه (7) دنانير، وأجرة العامل (25) دينار، فإن الإقتران الذي يمثل تكاليف سياج الحديقة هو:

(أ) 4س+25 　　(ب) 7س+25 　　(ج*) 28س+25 　　(د) 28س ـ 25

96) ينتج مصنع للثلاجات (س) ثلاجة يومياً، تكلفتها الكلية تساوي (20س + 25)، ويبيع الثلاجة الواحدة بمبلغ (375) دينار، فإن الربح (ص) الذي يحققه المصنع يومياً بدلالة عدد الثلاجات المنتجة هو:

(أ) ص= 375+(20س+25) 　　(ب) ص= (20س+25) ـ 375

(ج) ص= 375 ـ 20س 　　(د*) ص = 355س ـ 25

انتهت الأسئلة

ملحق (5)
اسماء وتخصصات وخبرات المحكمين ودرجاتهم الأكاديمية

الرقم	الاسم	المؤهل العلمي	الرتبة الأكاديمية	الوظيفة	سنوات الخبرة
1	إيمان رسمي عبد	دكتوراه، مناهج وطرق تدريس الرياضيات	استاذ	عميد البحث العلمي (كلية العلوم التربوية وآدابها/الأنروا جامعة ناعور)	4
2	اسماعيل علي صالح	ماجستير، اساليب تدريس الرياضيات	-	مشرف تربوي (مديرية قصبة عمان)	24
3	خالد ابراهيم خضر	ماجستير، رياضيات بكا رياضيات	-	مشرف تربوي (مديرية لواء القويسمة)	24
4	هلا الشوا	دكتوراه، مناهج وطرائق تدريس الرياضيات	استاذ مشارك	عضو هيئة تدريس (الجامعة الأردنية)	8
5	خالد أبو لوم	دكتوراه، اساليب تدريس رياضيات	استاذ مشارك	عضو هيئة تدريس (الجامعة الأردنية)	20
6	حيدرابراهيم ظاظا	دكتوراه، قياس وتقويم	استاذ مشارك	عضو هيئة تدريس (الجامعة الأردنية)	10
7	فريال محمد أبو عوَاد	دكتوراه، قياس وتقويم	استاذ مشارك	عضو هيئة تدريس (الجامعة الأردنية)	7
8	جهاد العناتي	دكتوراه، قياس وتقويم	استاذ مساعد	عضو هيئة تدريس (الجامعة الأردنية)	4
9	خليل زهدي عبد الرزاق	دكتوراه، مناهج وطرق تدريس الرياضيات	-	مشرف تربوي (مديرية قصبة عمان)	28
10	أحمد القضاة	دكتوراه، أساليب تدريس الرياضيات	استاذ مشارك	عضو هيئة تدريس (جامعة آل البيت)	13
11	خميس نجم	دكتوراه، مناهج الرياضيات وأساليب تدريسها	استاذ	عضو هيئة تدريس (جامعة آل البيت)	11
12	ابراهيم أحمد الشرع	دكتوراه، اساليب تدريس الرياضيات	استاذ	رئيس قسم المناهج وأساليب تدريس(الجامعة الاردنية)	10
13	موسى عزت عوده	دكتوراه، قياس وتقويم، بكا رياضيات	-	مشرف تربوي (مديرية لواء القويسمة)	25

ملحق (6)

استبانة التحكيم (آراء المحكمين والنموذج المرفق مع الاختبار)

عزيزي المحكّم المحترم،

السلام عليكم ورحمة الله وبركاته،

أضع بين يديك اختبار تحصيلي في مادة الرياضيات (الجزء الأول) للصف الثامن الأساسي بهدف قياس تحصيل الطلبة لغايات بحثية. ويتكون هذا الاختبار من مجموعة من الفقرات (الأسئلة) من نوع الاختيار من متعدّد، حيث تمّ كتابة مجموعة من الأهداف المعرفية مصنّفة وفق تصنيف بلوم للأهداف المعرفية، يلي كل منها عدد متفاوت من الفقرات التي تقيس كل هدف، علماً بأنه تمّت الإشارة للإجابة الصحيحة باستخدام الرمز (*) بمحاذاة البديل الذي يمثّل الإجابة الصحيحة. أرجو التكرم بالاطلاع والتدقيق بعناية والنظر بعين خبيرة وناقدة للأهداف والفقرات التي تقيسها، والحكم على كل فقرة من فقرات الاختبار من خلال وضع إشارة (√) أمام البديل المناسب في النموذج المرفق في نهاية الاختبار من حيث:

1) وضوح وسلامة الصياغة اللغوية (واضحة، غير واضحة).
2) سلامة المحتوى العلمي (سليمة، غير سليمة).
3) درجة توافق الفقرة مع الهدف الذي تقيسه (كبيرة، متوسطة، ضعيفة، لا أدري).
4) ملاءمة مموهات الإجابة (ملائمة، غير ملائمة).
5) إضافة أية ملاحظات أو مقترحات أو تعديلات ترونها مناسبة.

اسم المحكّم	
المؤهل العلمي والرتبة الأكاديمية	
الوظيفة	
سنوات الخبرة	

الفقرة	**وضوح الصياغة اللغوية**		**سلامة المحتوى العلمي للفقرة**		**درجة توافق الفقرة مع الهدف الذي تقيسه**				**ملاءمة مموهات الاجابة**		**ملاحظات**
	واضحة	**غير واضحة**	**سليمة**	**غير سليمة**	**كبيرة**	**متوسطة**	**ضعيفة**	**لا أدري**	**ملاءمة**	**غير ملاءمة**	
1											
2											
3											
4											

ملحق (7)
نتائج التحكيم (عدد المحكمين المتفقين وغير المتفقين على الفقرات)

رقم الفقرة	وضوح الصياغة اللغوية		سلامة المحتوى العلمي للفقرة		درجة توافق الفقرة مع الهدف الذي تقيسه			ملاءمة موهات الاجابة	
	واضحة	غير واضحة	سليمة	غير سليمة	كبيرة	متوسطة	ضعيفة	ملاءمة	غير ملاءمة
1	13		13		13			13	
2	13		13		12	1		13	
3	12	1	13		12	1		13	
4	13		13		13			13	
5	13		13		13			11	2
6	12	1	13		13			13	
7	13		13		13			13	
8	13		13		10	2	1	13	
9	13		13		13			13	
10	**9**	**4**	**8**	**5**	**8**	**3**	**2**	**10**	**3**
11	13		12	1	13			13	
12	13		13		13			12	1
13	13		13		13			11	2
14	**9**	**4**	**9**	**4**	**8**	**3**	**2**	**13**	
15	12	1	13		13			12	1
16	13		13		12	1		13	
17	13		13		13			10	3
18	13		13		13			10	3
19	**9**	**4**	**9**	**4**	**9**	**2**	**2**	**9**	**4**
20	12	1	13		13			11	2
21	13		13		13			13	
22	13		13		12	1		12	1
23	13		13		11	2		12	1
24	11	2	13		13			13	
25	13		13		12	1		13	
26	13		13		11	1	1	13	
27	12	1	13		13			13	
28	13		13		10	2	1	11	2
29	13		13		13			13	
30	13		13		13			11	2
31	13		13		13			13	
32	13		13		13			13	

رقم الفقرة	وضوح الصياغة اللغوية		سلامة المحتوى العلمي للفقرة		درجة توافق الفقرة مع الهدف الذي تقيسه			ملاءمة مموهات الاجابة	
	واضحة	غير واضحة	سليمة	غير سليمة	كبيرة	متوسطة	ضعيفة	ملاءمة	غير ملاءمة
33	12	1	13		13			13	
34	13		13		12	1		13	
35	13		13		13			13	
36	12	1	13		13			13	
37	13		13		13			13	
38	13		13		13			13	
39	**8**	**5**	**10**	**3**	**12**	**1**		**13**	
40	**7**	**6**	**9**	**4**	**11**	**1**	**1**	**12**	**1**
41	12	1	13		13			13	
42	13		13		13			13	
43	13		13		13			12	1
44	12	1	13		13			13	
45	13		13		13			13	
46	13		13		13			13	
47	13		11	2	12	1		12	1
48	13		13		13			13	
49	13		13		13			11	2
50	13		13		13			13	
51	13		13		13			11	2
52	13		13		12	1		13	
53	13		13		13			13	
54	13		13		13			13	
55	12	1	13		13			13	
56	13		13		13			13	
57	13		12	1	13			13	
58	13		13		13			13	
59	13		13		13			13	
60	13		13		13			13	
61	13		13		13			13	
62	12	1	13		12	1		13	
63	13		13		13			13	
64	13		13		13			11	2
65	13		13		13			13	
66	13		13		13			13	
67	13		13		10	2		12	1

رقم الفقرة	وضوح الصياغة اللغوية		سلامة المحتوى العلمي للفقرة		درجة توافق الفقرة مع الهدف الذي تقيسه			ملاءمة مموهات الاجابة	
	واضحة	غير واضحة	سليمة	غير سليمة	كبيرة	متوسطة	ضعيفة	ملاءمة	غير ملاءمة
68	13		13		13			13	
69	13		13		13			13	
70	13		13		13			13	
71	13		13		13			13	
72	13		13		13			13	
73	13		13		13			13	
74	13		13		11	1	1	11	2
75	13		13		13			13	
76	13		13		13			13	
77	13		13		11	2		13	
78	12	1	13		13			13	
79	13		13		13			12	1
80	13		13		13			13	
81	13		13		13			13	
82	13		13		13			13	
83	13		13		13			13	
84	13		13		11	2		12	1
85	13		13		13			13	
86	13		13		13			13	
87	13		13		13			13	
88	12	1	13		13			13	
89	13		12	1	11	2		12	1
90	13		13		13			13	
91	13		13		13			13	
92	13		13		11	2		13	
93	13		13		13			13	
94	13		13		13			13	
95	13		13		13			13	
96	12	1	13		12	1		13	

ملحق (8)
كتاب تسهيل المهمة

وزارة التربية والتعليم
مديرية التربية والتعليم للواء قصبة عمان

الرقم ٢٩٥٨/١/٧٣
التاريخ ٢٠/٥/١٤٣٧هـ
الموافق ٢٨/٢/٢٠١٦م

مديري ومديرات المدارس

الموضوع / تسهيل مهمة

السـلام عليكم ورحمـة الله وبركاتـه

إشـارة لكتـاب رئيس الجامعة الأردنية رقم ١/٢٠١٦/١١٩ الموافق ١١/١/٢٠١٦م.
يقوم الطالب " نبيل موسى جبران جبران من طلبة برنامج دكتوراه علم النفس التربوي القياس والتقويم في كلية العلوم التربوية بالجامعة الأردنية " بإعداد أطروحة عنوانها:

" أثر حجم العينة وعدد الفقرات وصور الاختبار على دقة تقدير معالم الفقرة والقدرة ودالة المعلومات لاختبار تحصيلي في الرياضيات للصف الثامن "
الأمر الذي يحتاج إلى تطبيق أداة دراستها على طلبة الصف الثامن الأساسي (ذكوراً وإناثاً) في مدارسكم.

آملاً تسهيل مهمة الطالب المذكور وتقديم المساعدة الممكنة له،على أن يتم مطابقة الاستبانة المرفقة مع الاستبانة المطبقة.

و تفضلوا بقبول الاحترام

/ مدير التربيـة والتعليم

سليمان النوافعة
مدير الشؤون التعليمية والفنية

- نسخة : مدير الشؤون التعليمية والفنية .
- نسخة : ر.ق التدريب والتأهيل والإشراف التربوي .
- نسخة : عضو قسـم الإشـراف .
- المرفقات: الاستبانة.

تلفون :(٦ – ٥٦٩٩١٨١ -٠٦) فاكس : (٥٦٩٩٥٨٠-٠٦) ص.ب : (٩٥٧٩ اللويبدة)

المملكة الأردنية الهاشمية
هاتف:٥٦٠٧١٨١ ٦ ٩٦٢+ فاكس: ٥٦٦٦٠١٩ ٦ ٩٦٢+ص.ب ١٦٤٦ عمان ١١١١٨ الأردن. الموقع الإلكتروني:www.moe.gov.jo

ملحق (9)
فاعلية المموهات للإختبار المكوَن من (96) فقرة

الفقرة	الإجابة	فاعلية المموهات				الفقرة	الإجابة	فاعلية المموهات			
		أ	ب	جـ	د			أ	ب	جـ	د
1	جـ	-0.14	-0.09	0.36	-0.14	28	د	0.05	-0.25	-0.18	0.39
2	ب	-0.27	0.39	-0.05	-0.07	29	أ	0.55	-0.25	-0.16	-0.14
3	د	-0.05	-0.09	-0.14	0.27	30	جـ	-0.09	-0.25	0.59	-0.25
4	ب	-0.20	0.45	-0.07	-0.18	31	أ	0.59	-0.07	-0.43	-0.09
5	أ	-0.05	0.11	-0.02	-0.05	32	جـ	-0.11	-0.09	0.36	-0.16
6	جـ	-0.20	-0.18	0.48	-0.09	33	ب	0.09	0.14	-0.02	-0.20
7	ب	0.00	0.32	-0.16	-0.16	34	أ	0.64	-0.32	-0.23	-0.09
8	أ	0.59	-0.05	-0.25	-0.30	35	جـ	-0.18	-0.30	0.70	-0.23
9	د	-0.07	-0.23	-0.25	0.55	36	ب	0.02	0.05	-0.07	0.00
10	**حذفت من قبل المحكمين**					37	د	-0.20	-0.14	-0.07	0.41
11	جـ	-0.07	-0.43	0.61	-0.11	38	ب	-0.34	0.57	-0.16	-0.07
12	د	-0.09	-0.05	-0.11	0.25	**39**	**حذفت من قبل المحكمين**				
13	أ	0.50	-0.23	-0.14	-0.14	**40**	**حذفت من قبل المحكمين**				
14	**حذفت من قبل المحكمين**					41	أ	0.45	-0.34	-0.07	-0.05
15	ب	-0.05	-0.05	-0.16	0.25	42	جـ	-0.05	-0.11	0.39	-0.23
16	أ	0.68	-0.16	-0.45	-0.07	43	د	-0.09	-0.07	-0.39	0.55
17	جـ	-0.23	-0.11	0.32	0.02	44	جـ	-0.20	-0.11	-0.23	-0.25
18	ب	-0.11	0.32	-0.14	-0.07	45	د	-0.23	-0.23	-0.05	0.50
19	**حذفت من قبل المحكمين**					46	أ	0.48	-0.02	-0.18	-0.27
20	أ	0.55	-0.20	-0.09	-0.25	47	جـ	-0.07	0.09	0.11	-0.14
21	جـ	-0.20	-0.05	0.30	-0.05	48	جـ	-0.16	0.25	0.02	-0.11
22	ب	-0.05	0.41	-0.11	-0.25	49	أ	0.25	-0.23	0.00	-0.02
23	جـ	-0.27	0.07	0.36	-0.16	50	د	-0.18	-0.09	-0.16	0.43
24	د	-0.05	-0.07	-0.09	0.20	51	ب	-0.11	0.52	-0.25	-0.16
25	ب	-0.27	0.52	-0.14	-0.11	52	د	-0.20	-0.45	-0.07	0.73
26	جـ	-0.14	-0.18	0.52	-0.20	53	أ	0.59	-0.11	-0.20	-0.27
27	جـ	-0.27	0.00	0.30	-0.02	54	ب	-0.09	0.36	-0.18	-0.09

الفقرة	الإجابة	فاعلية المموهات			
		أ	ب	جـ	د
55	جـ	-0.27	-0.14	0.45	-0.05
56	ب	-0.25	0.39	-0.09	-0.05
57	جـ	-0.09	-0.18	0.34	-0.07
58	د	-0.16	-0.05	0.02	0.18
59	أ	0.52	-0.18	-0.16	-0.18
60	جـ	-0.39	-0.11	0.61	-0.11
61	د	-0.25	-0.07	-0.07	0.39
62	أ	0.59	-0.36	-0.09	-0.14
63	ب	-0.27	0.64	-0.39	0.02
64	جـ	-0.05	-0.18	0.39	-0.16
65	د	-0.09	-0.18	-0.23	0.50
66	أ	0.57	-0.27	-0.16	-0.14
67	د	-0.16	-0.11	-0.11	0.39
68	أ	0.45	-0.23	-0.20	-0.02
69	ب	-0.05	0.41	-0.11	0.14
70	أ	0.09	-0.02	-0.02	-0.05
71	جـ	-0.02	-0.02	0.09	-0.05
72	ب	-0.02	0.27	-0.16	-0.09
73	د	-0.14	-0.27	-0.11	0.52
74	جـ	-0.09	0.02	0.25	-0.18
75	أ	0.52	-0.09	-0.27	-0.16
76	ب	-0.25	0.45	-0.11	-0.09
77	جـ	-0.25	-0.23	0.43	0.05
78	أ	0.43	-0.07	-0.25	-0.11
79	د	-0.34	-0.25	-0.11	0.70
80	ب	-0.27	0.55	-0.16	-0.11
81	أ	0.52	-0.25	-0.16	-0.11
82	جـ	0.02	-0.14	0.25	-0.14
83	أ	0.30	-0.25	0.00	-0.05
84	د	-0.16	-0.25	-0.05	0.45
85	ب	-0.11	0.41	-0.27	-0.02
86	جـ	-0.16	-0.27	0.48	-0.05
87	أ	0.39	-0.25	-0.11	-0.02
88	د	-0.09	-0.14	-0.36	0.59
89	ب	-0.14	0.45	-0.16	-0.16
90	جـ	-0.05	-0.07	0.27	-0.16
91	ب	0.23	-0.09	0.00	-0.14
92	أ	0.48	-0.16	-0.18	-0.14
93	جـ	-0.09	-0.20	0.43	-0.14
94	ب	-0.18	0.36	-0.09	-0.09
95	جـ	-0.05	0.11	0.05	-0.11
96	د	0.07	0.16	-0.11	-0.11

ملحق(10)
أرقام فقرات الاختبار في صورته الأولية وأرقامها الجديدة في الاختبارين (50، 25 فقرة)

رقم الفقرة في الاختبار بصورته الأولية(96فقرة)	الرقم للفقرة في اختبار (50 فقرة)	الرقم للفقرة في اختبار (25 فقرة)	رقم الفقرة في الاختبار بصورته الأولية(96فقرة)	الرقم للفقرة في اختبار (50 فقرة)	الرقم للفقرة في اختبار (25 فقرة)
1	1		25	10	
2		1	26		5
3	4		27		6
4	2		28		7
5	حذفت لتدني معامل التمييز		29	فقرة زائدة	
6	3		30	11	
7		2	31	12	
8	21		32		8
9			33	حذفت لتدني معامل التمييز	
10	حذفت من قبل المحكمين		34	13	
11		9	35	14	
12		13	36	حذفت لتدني معامل التمييز	
13		4	37	15	
14	حذفت من قبل المحكمين		38	فقرة زائدة	
15	حذفت لتدني معامل تمييزها		39	حذفت من قبل المحكمين	
16	35		40	حذفت من قبل المحكمين	
17		25	41	16	
18	5		42	33	
19	حذفت من قبل المحكمين		43	17	
20	6		44	18	
21	7		45	19	
22	8		46	20	
23		10	47	حذفت لتدني معامل التمييز	
24	9		48	حذفت لتدني معامل التمييز	

رقم الفقرة في الاختبار بصورته الأولية	الرقم للفقرة في اختبار (50 فقرة)	الرقم للفقرة في اختبار (25 فقرة)	رقم الفقرة في الاختبار بصورته الأولية	الرقم للفقرة في اختبار (50 فقرة)	الرقم للفقرة في اختبار (25 فقرة)
49		11	73	36	
50	22		74	37	
51	23		75	39	
52	24		76		19
53	25		77	40	
54	26		78	41	
55	27		79	42	
56	29		80	43	
57		12	81	44	
58		24	82	38	
59	28		83		20
60	حذفت لتدني معامل التمييز		84	45	
61		14	85	46	
62		15	86		21
63		16	87	حذفت لتدني معامل الصعوبة	
64	حذفت لتدني معامل الصعوبة		88	47	
65	30		89		22
66		17	90	48	
67	31		91	حذفت لتدني معامل التمييز	
68	32		92		23
69	18		93	49	
70	حذفت لتدني معامل التمييز		94	50	
71	34		95	حذفت لتدني معاملي التمييز والصعوبة معا.	
72	حذفت لتدني معامل التمييز		96	حذفت لتدني معاملي التمييز والصعوبة معا.	

ملحق (11)

جدول مواصفات للاختبارين (طويل 50 فقرة، و قصير 25 فقرة)

أولاً: جدول مواصفات الإختبار (50 فقرة):

المحتوى	عدد الصفحات	النسبة تقريباً	مستوى الأهداف (النسبة المؤية)			المجموع
			معرفة وتذكر (%30)	الفهم والاستيعاب (%40)	التطبيق (%30)	
الأعداد الحقيقية	63	43%	6	9	6	21
المقادير الجبرية	37	26%	4	5	4	13
الأنماط والإقترانات	45	32%	5	6	5	16
المجموع	145	100%	15	20	15	50

ثانياً: جدول مواصفات الإختبار (25 فقرة):

المحتوى	عدد الصفحات	النسبة تقريباً	مستوى الأهداف (النسبة المؤية)			المجموع
			معرفة وتذكر (%30)	الفهم والاستيعاب (%40)	التطبيق (%30)	
الأعداد الحقيقية	63	43%	3	5	3	11
المقادير الجبرية	37	26%	2	3	2	7
الأنماط والإقترانات	45	32%	2	3	2	7
المجموع	145	100%	7	11	7	25

ملحق (12)

اختبار طويل (50 فقرة) من نوع الإختيار من متعدد في صورته النهائية بعد التحكيم

عزيزي الطالب/ الطالبة:....................................... يتكون هذا الاختبار من (50) فقرة من نوع الاختيار من متعدد حيث لكل فقرة أربعة بدائل واحد منها فقط صحيح، اقرأجميع الأسئلة ثم ضع دائرة حول رمز الاجابة الصحيحة.

---

1) واحد من الاعداد الآتية يعد عدد <u>غير</u> نسبي:

(أ) 8 (ب) – 8 (جـ) $\sqrt{8}$ (د) $0,\overline{8}$

2) أحد الأعداد الآتية <u>لا</u> يمكن كتابته على الصورة $\frac{أ}{ب}$ ، ب≠0:

(أ) 0,65 (ب) $\sqrt{23}$ (جـ) 23 (د) $0,2\overline{3}$

3) قيمة المقدار $\sqrt[3]{-64}$ تساوي:

(أ) -8 (ب) 8 (جـ) -4 (د) 4

4) العدد ... 0,151151115 هو عدد:

(أ) صحيح (ب) طبيعي (جـ) نسبي (د) غير نسبي

5) إذا كانت س، ص، ع ثلاثة أعداد حقيقية وكان (س+ص)+ع = س+(ص+ع)، فإن ذلك يعني أن الأعداد الحقيقية تحقق الخاصية:

(أ) التبديلية (ب) التجميعية (جـ) توزيع الجمع على الضرب (د) توزيع الضرب على الجمع

6) العنصر المحايد لعملية الجمع على مجموعة الاعداد الحقيقية هو:

(أ) صفر (ب) 1 (جـ) -1 (د) غير موجود

7) النظير الضربي للعدد $\frac{3}{4}$ هو:

(أ) $\frac{-3}{4}$ (ب) 1 (جـ) $\frac{4}{3}$ (د) صفر

8) قيمة المقدار $\sqrt{3}$ ($\sqrt{3}$ + $\sqrt{12}$) تساوي:

(أ) 3 (ب) 9 (جـ) $3\sqrt{12}$ (د) $3\sqrt{15}$

9) احدى الفترات الاتية فترة مغلقة:

(أ) (1، 5) (ب) [1، 5) (ج) (1 ، 5] (د) [1 ، 5]

10) تعبر المجموعة : ف ={ س : س ∍ ح، -4< س ≤ 3 } عن الفترة:

(أ) [-4 ، 3] (ب) (-4 ، 3] (ج) [-4 ، 3) (د) (-4 ،3)

11) مجموعة حل المتباينة : س- 7 > 5 هي:

(أ) (-12 ، ∞) (ب)] 12 ، ∞] (ج) (12 ، ∞) (د) (- ∞ ، 12)

12) مجموعة حل المتباينة $\frac{3}{4}$ س ≥ $\frac{7}{8}$ هو:

(أ) (-∞ ، $\frac{7}{6}$] (ب) (-∞ ، $\frac{7}{6}$) (ج) ($\frac{3}{4}$ ، $\frac{7}{8}$] د) (-∞ ، ∞)

13) قيمة المقدار (($\sqrt{5}$)^3)^2 تساوي

(أ) 125 (ب) 225 (ج) 25 (د) 5

14) قيمة المقدار ($\sqrt{12}$)$^{-4}$ تساوي

(أ) 144 (ب) $\sqrt{12}$ (ج) $\frac{1}{144}$ (د) - $\frac{1}{144}$

15)) يكتب العدد (4083000000) على الصورة العلمية على النحو:

(أ) 4,083 × 10^{-9} (ب) 408 ,3 ×10^{9} (ج) 40,83 × 10^{9} د) 4,083×10^{9}

16) $(64)^{3/1}$ =

(أ)4 (ب)8 (جـ) -4 (د) -8

17) $(\sqrt{17})^{4/3} \times (\sqrt{17})^{4/5} =$

(أ) -17 (ب) $\sqrt{17}$ (ج) $\sqrt[3]{17}$ (د) 17

18) يكتب العدد $\sqrt[5]{\frac{-243}{32}}$ في أبسط صورة على النحو:

(أ) $\frac{3}{2}$ (ب) $\frac{2}{3}$ (ج) $\frac{-3}{2}$ د) $-\frac{2}{3}$

19) عند اختصار المقدار $\frac{س^{4/3} \times ص^{3/2}}{س^{2/1} \times ص^{6/1}}$ الى أبسط صورة يصبح:

(أ) $س^{6/4}ص^{9/5}$ (ب) س $ص^{-3/1}$ (ج) $س^{2/1}ص^{4/1}$ (د) $س^{4/1}ص^{2/1}$

20) قيمة المقدار $(\sqrt{3} \times \sqrt[3]{9})^6$ تساوي:

(أ) 2187 (ب) 729 (ج) 243 (د) 180

21) يمكن كتابة العدد الكسري $2\frac{7}{9}$ على الصورة :

(أ) $(\frac{5}{3})^2$ (ب) $(\frac{3}{5})^2$ (ج) $(\frac{7}{9})^2$ (د) $(2\frac{7}{9})^2$

22) القسم الرمزي للحد الجبري ($\sqrt{5}$ $س^3$ص ع) هو:

(أ) $\sqrt{5}$ (ب) $س^3$ (ج) $س^3$ص (د) $س^3$ص ع

23) معامل الحد الجبري ($\frac{5}{7}$ س ص) هو:

(أ) $\frac{5}{7}$ ص (ب) $\frac{5}{7}$ (ج) س ص (د) $\frac{5}{7}$ س

24) حديقة على شكل مستطيل طولها (س) متراً وعرضها (ص) متراً فإن المقدار الجبري الذي يمثل محيطها

(أ) س ص (ب) س+ص (ج) 2س +ص (د) 2س + 2ص

- **اذا كانت س = 5، ص = ـ3، ع = 6، اجب عن السؤالين (25،26)، الآتيين:**

25) القيمة العددية للمقدار الجبري (7س ص) تساوي :

(أ) – 105　　(ب) 105　　(ج) 35　　(د) -21

26) القيمة العددية للمقدار (2س ع + ص^2) تساوي:

أ) 51　　(ب)69　　(ج) -30　　(د) 60

27) الحد الجبري الذي يشابه الحد الجبري (7س^{3}ع) هو:

(أ) 7س ع　　(ب) س ع^3　　(ج) 8س^{3}ع　　(د) 6س^{3}ع^3

28) 4ص^2 ع^3 ×3ع^2 ص^3 =

(أ)12ص^5 ع^5　　(ب)12ص^2 ع^5　　(ج)12ص^{5}ع^3　　(د)12ص^4 ع^6

29) 2س + 5س يساوي:

(أ) 7س^2　　(ب) 7س　　(ج) 7　　(د) س^2

30) العامل المشترك الاكبر(ع.م.أ) للحدين (3س^2 ، 6س)هو:

(أ) 2س　　(ب) 18س^3　　(ج) 9س^3　　(د) 3س

31) يكتب المقدار الجبري (8أ^{2}ب – 10أ ب^2) على الصورة:

(أ) 2أب(4أ+5ب)　　(ب) 2أب(أ^2 – ب^2)　　(ج) 2أب(ب^2- أ^2)　(د) 2أب(4أ- 5ب)

32) يكتب المقدار الجبري (8س^{2}ص^3+18ص^{2}س^3) على الصورة:

(أ) 2س^{2}ص^2(4ص+9س)　　(ب)8س^{2}ص^2(ص+8س)

(جـ)2س^{2}ص^2(ص^3+18س)　　(د)2س ص(4ص+9س)

33) قيمة المقدار (-128)$^{7/1}$ تساوي:

(أ) 2　　(ب) 3　　(جـ) -2　　(د) -3

34) قطعة أرض مستطيلة الشكل عرضها (2س) متر، وطولها ثلاثة أمثال عرضها فإن مساحتها بدلالة س هي:

(أ) 12س　　(ب) 16س^2　　(ج)12س^2　　(د) 6س^2

35) $\sqrt{3} + 5\sqrt{3}$ يساوي :

(أ) $6\sqrt{3}$ (ب) $5\sqrt{9}$ (جـ) $5\sqrt{6}$ (د) $6\sqrt{6}$

36) يحتوي خزان على 625م^3 من الماء، وينقص حجم الماء كل يوم بمقدار 25م^3 عن اليوم الذي قبله، اذا كان
عدد الأيام = س فإن العلاقة التي تحدد كمية الماء في الخزان بدلالة عدد الأيام هي:

(أ) 625+25س (ب) 625×25س (ج) 625÷25س (د) 625 – 25س

37) احدى من العبارات الآتية صحيحة:

(أ) كل علاقة اقتران (ب) بعض الاقترانات علاقات
(جـ) كل اقتران علاقة (د) بعض الاقترانات ليست علاقات

38) أي من الاقترانات الاتية اقتران ثابت:

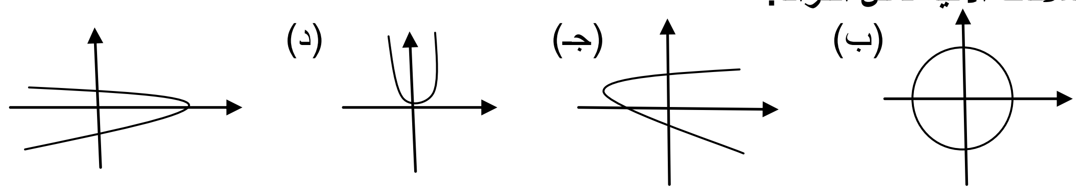
(أ) ق(س)= 5س+2 (ب) ل(س)= 8- س (ج) م(س)= $\sqrt{3}$ (د) هـ(س)= 4+2س

39) في الزوج المرتب (س،ص)، فإن س تمثل:

(أ) المجال (ب) المدى (ج) قاعدة الاقتران (د) بيان الاقتران

40) أي العلاقات الآتية تمثل اقتراناً:

(أ) (ب) (جـ) (د)



41) احدى العلاقات الآتية تمثل اقتران:

(أ) {(-1،5)،(0،2)(3،1)} (ب) {(2،1)،(5،1)،(4،2)}
(جـ) {(3،3)،(1،2)،(0،3)} (د) {(1،2)،(5،3)،(4،2)}

42) قاعدة الاقتران ق(س) الذي يمثله الجدول الاتي هي:

س	5	7	9	11
ص	9	13	17	21

(أ) ق(س)=2س (ب) ق(س)= س^2 (ج) ق(س)=2س+1 (د)ق(س)= 2س -1

43) أي مما يأتي يمثل اقتران خطي:

(أ) ق(س)= $\sqrt{س}$ +3　　(ب) ل(س)=2س-3　　(ج)هـ(س)=4س$^{-2}$-3　　(د) م(س)=5س^{2}+3

44) الصورة العامة للإقتران الخطي هي:

(أ) ق(س)=أس+ب　　(ب) ل(س)= أس×ب س　　(ج) هـ(س)= أس^{2}+ب　　(د) ك(س)= $\frac{1}{س}$ + ب

45) اذا كان ق(س)=3س-2 فإن ق($\frac{1}{3}$) يساوي:

(أ) 7　　(ب) 1　　(جـ) -2　　(د) -1

46) ميل منحنى الاقتران ق(س) = 7-4س هو:

(أ) 4　　(ب) -4　　(جـ) -7　　(د) 7

47) منحنى الاقتران هـ(س)= 6-3س يكون:

(أ) ثابتاً　　(ب) متزايداً　　(ج) موازياً لمحور السينات　　(د) متناقصاً

48) ميل الاقتران ق(س)=12 يساوي:

(أ) 12　　(ب) 1　　(ج) صفر　　(د) لاميل له

49) أراد عامل اكمال بناء سور من الحجر، فوجد انه تم بناء (95) حجر قبل أن يبدأ بالعمل، فإذا قام العامل ببناء (23) حجر يومياً حتى أكمل بناء السورخلال (س) من الأيام فإن المقدار الجبري الذي يعبر عن عدد الحجارة التي يتكون منها السور هو:

(أ) 95-3س　　(ب) 23س - 95　　(ج) 95+23س　　(د) 95×23س

50) إذا كان ثمن تذكرة دخول مدينة ألعاب ترفيهية (دينار) للفرد الواحد، وأجرة الاشتراك في أي لعبة من الألعاب (50 قرشاً) اشترك أحمد في (6) ألعاب، المبلغ الذي دفعه بالدنانير هو:

(أ) 3　　(ب) 4　　(جـ) 6　　(د) 7

انتهت الأسئلة

ملحق (13)

اختبار طويل (50 فقرة) من نوع الصواب والخطأ

اختبار تحصيلي في مادة الرياضيات للصف الثامن الأساسي (الفصل الأول)   مدة الاختبار: 45 دقيقة

عزيزي الطالب/ الطالبة:....................................... يتكون هذا الاختبار من (50) فقرة من نوع الصواب والخطأ، اقرأجميع الأسئلة ثم ضع اشارة (√) أمام العبارة الصحيحة واشارة (x) أمام العبارة الخطأ.

الأسئلة :

1) العدد $\sqrt{8}$ عدد غير نسبي.

2) العدد $\sqrt{23}$ لا يمكن كتابته على صورة $\frac{أ}{ب}$

3) قيمة المقدار $\sqrt[3]{-64}$ = 4

4) العدد ... 0,151151115 هو عدد: غير نسبي .

5) إذا كانت س، ص، ع ثلاثة أعداد حقيقية وكان (س+ص)+ع = س+(ص+ع)، فإن ذلك يعني أن الأعداد الحقيقية تحقق خاصية توزيع الضرب على الجمع.

6) العنصر المحايد لعملية الجمع على مجموعة الاعداد الحقيقية هوالعدد صفر.

7) النظير الضربي للعدد $\frac{3}{4}$ هو العدد $\frac{-3}{4}$.

8) قيمة المقدار $\sqrt{3}$ ×($\sqrt{3}$ + $\sqrt{12}$) تساوي 9.

9) الفترة [1 ، 5] هي فترة مغلقة.

10) تعبر المجموعة : ف ={ س : س ∍ ح، -4< س ≤ 3 } عن الفترة   [-4 ، 3).

11) مجموعة حل المتباينة : س- 7 > 5 هي: (12 ، ∞).

12) مجموعة حل المتباينة $\frac{3}{4}$ س ≥ $\frac{7}{8}$ هو ($\frac{3}{4}$ ، $\frac{7}{8}$].

13) قيمة المقدار (($\sqrt{5}$)^3)^2 تساوي 125.

14) قيمة المقدار ($\sqrt{12}$)$^{-4}$ تساوي - $\frac{1}{144}$.

15) يكتب العدد (4083000000) على الصورة العلمية على الشكل ($4,083\times10^9$).

16) قيمة المقدار (64)$^{3/1}$ تساوي 8

17) ($\sqrt{17}$)$^{4/3}$× ($\sqrt{17}$)$^{4/5}$ = $\sqrt{17}$

18) يكتب العدد $\sqrt[5]{\frac{-243}{32}}$ في أبسط صورة على الشكل – $\frac{2}{3}$

19) عند اختصار المقدار $\frac{\text{س}^{4/3}\times\text{ص}^{3/2}}{\text{س}^{2/1}\times\text{ص}^{6/1}}$ الى أبسط صورة يصبح س$^{4/1}$ص$^{2/1}$

20) قيمة المقدار ($\sqrt{3}$ × $\sqrt[3]{9}$)^6 تساوي 729

21) يمكن كتابة العدد الكسري $2\frac{7}{9}$ على الصورة : ($\frac{5}{3}$)^2.

22) القسم الرمزي للحد الجبري ($\sqrt{5}$ س^{3}ص ع) هو س^{3}ص ع.

23) معامل الحد الجبري ($\frac{5}{7}$ س ص) هو $\frac{5}{7}$ س.

24) حديقة على شكل مستطيل طولها (س) متراً وعرضها (ص) متراً فإن المقدار الجبري الذي يمثل محيطها هو س + ص .

- **اذا كانت س = 5، ص = ـ3، ع = 6، اجب عن السؤالين (25،26)، الآتيين:**

25) القيمة العددية للمقدار الجبري (7س ص) تساوي – 105

26) القيمة العددية للمقدار (2س ع + ص^2) تساوي 51

27) الحد الجبري الذي يشابه الحد الجبري (7س^{3}ع) هو 8س^{3}ع

28) 4 ص^2 ع^3 ×3ع^2 ص^3 = 12ص^5 ع^5

29) 2س + 5س يساوي: **7س^2**

30) العامل المشترك الاكبر(ع.م.أ) للحدين (3س^2 ، 6س)هو 3س

31) يكتب المقدار الجبري (8أ^{2}ب – 10أ ب^2) على الصورة 2أب(4أ- 5ب)

32) يكتب المقدار الجبري (8س^{2}ص^3+18ص^{2}س^3) على الصورة 2س^{2}ص^2(4ص+9س)

33) قيمة المقدار (-128)$^{7/1}$ يساوي: **-2**

34) قطعة أرض مستطيلة الشكل عرضها (2س) متر، وطولها ثلاثة أمثال عرضها فإن مساحتها بدلالة س هي (6س^2) .

35) $\sqrt[3]{\ }$ +5$\sqrt[3]{\ }$ يساوي : 6$\sqrt[3]{\ }$

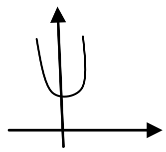
36) يحتوي خزان على 625م^3 من الماء، وينقص حجم الماء كل يوم بمقدار 25م^3 عن اليوم الذي قبله، اذا كان

عدد الأيام = س، فإن العلاقة التي تحدد كمية الماء في الخزان بدلالة عدد الأيام هي: 625÷25س

37) بعض الاقترانات ليست علاقات .

38) الاقتران م(س)= $\sqrt{3}$ اقتران ثابت.

39) في الزوج المرتب (س،ص)، فإن س تمثل المدى.



40) العلاقة المرسومة في الشكل تمثل اقتراناً:

41) العلاقة {(-5،1)،(2،0)(1،3)} تمثل اقتران .

42) قاعدة الاقتران ق(س) الذي يمثله الجدول الاتي هي ق(س)= 2س +1

س	5	7	9	11
ص	9	13	17	21

43) الصيغة ق(س)= $\sqrt{س}$ +3 تمثل اقتران خطي .

44) الصورة العامة للإقتران الخطي هي ق(س)=أس+ب .

45) اذا كان ق(س)=3س-2 فإن ق($\frac{1}{3}$) يساوي -1

46) ميل منحنى الاقتران ق(س) = 7-4س هو -7

47) منحنى الاقتران هـ(س)= 6-3س يكون متزايداً .

48) ميل الاقتران ق(س)=12 يساوي 1

49) أراد عامل اكمال بناء سور من الحجر، فوجد انه تم بناء (95) حجر قبل أن يبدأ بالعمل، فإذا قام العامل ببناء (23) حجراً يومياً حتى أكمل بناء السورخلال (س) من الأيام فإن المقدار الجبري الذي يعبر عن عدد الحجارة التي يتكون منها السور هو 95+23س

50) إذا كان ثمن تذكرة دخول مدينة ألعاب ترفيهية (دينار) للفرد الواحد، وأجرة الاشتراك في أي لعبة من الألعاب (50 قرشاً) فإذا اشترك أحمد في (6) ألعاب،فإنه سيدفع مبلغ 6 دنانير.

انتهت الأسئلة

ملحق (14)

اختبار قصير (25 فقرة) من نوع الاختيار من متعدد في صورته النهائية بعد التحكيم

اختبار تحصيلي في الرياضيات للصف الثامن الاساسي (الفصل الاول) مدة الاختبار: 30 دقيقة

عزيزي الطالب/ الطالبة:...................................... يتكون هذا الاختبار من (25) فقرة من نوع الاختيار من متعدد حيث لكل فقرة أربعة بدائل واحد منها فقط صحيح، اقرأجميع الأسئلة ثم ضع دائرة حول رمز الاجابة الصحيحة.

---

1) إحدى العبارات الآتية صحيحة:

(أ) كل عدد حقيقي هو عدد نسبي (ب) كل عدد نسبي هو عدد حقيقي

(جـ) كل عدد حقيقي هو عدد غير نسبي (د) كل عدد نسبي هو عدد صحيح

2) قيمة المقدار $\sqrt{0,16}$ =

(أ) 4 (ب) 0,4 (جـ) 0,8 (د) 0,04

3) **قيمة المقدار $\sqrt[3]{\frac{27}{64}}$ يساوي:**

(أ) $\frac{9}{8}$ (ب) $\frac{3}{8}$ (جـ) $\frac{27}{64}$ (د) $\frac{3}{4}$

4) **قيمة المقدار $\sqrt[3]{125\times64}$ يساوي:**

(أ) 20 (ب) 40 (ج) 0,20 (د) 0,40

5) **عند ضرب طرفي المتباينة $\sqrt{3} < 2$ بالعدد (-1) تصبح:**

(أ) $-\sqrt{3} > 2$ (ب) $-\sqrt{3} > -2$ (ج) $-\sqrt{3} < -2$ (د) $-\sqrt{3} < 2$

6) **إحدى الفترات الاتية فترة غير محدودة:**

(أ) (2 ، 3) (ب) [2 ، 3] (ج) [2 ، ∞) (د) (2، 3]

7) **طول الفترة [-2 ، 3] يساوي:**

(أ) 1 (ب) -1 (ج) -5 (د) 5

8) **مجموعة حل المتباينة 2س + 7 > 9- 4س هو:**

(أ) $(-\infty ، \frac{1}{3})$ (ب) $(-\frac{1}{3} ، \frac{1}{3})$ (ج) $(\frac{1}{3} ، \infty)$ (د) $[\frac{1}{3} ، \infty)$

9) **قيمة المقدار $\sqrt[3]{15\frac{5}{8}}$ تساوي :**

(أ) $\frac{5}{8}$ (ب) $\frac{125}{8}$ (جـ) $\frac{5}{2}$ (د) $\frac{2}{5}$

10) **يكتب العدد $\frac{1}{\sqrt{10}+3}$ بحيث لا يظهر الجذر في المقام على الصورة:**

(أ) $\sqrt{10}+3$ (ب) $\frac{1}{3}$ (ج) $\sqrt{10}-3$ (د) $\frac{1}{\sqrt{10}}$

11) **مستطيل طوله $\sqrt{360}$ سم، وعرضه $\sqrt{40}$ سم، فإن مساحته بالسنتيمتر المربع بأبسط صورة تساوي:**

(أ) 120 (ب) 20 (ج) 14400 (د) 12

12) **(6س + 3ص) + (5ع – 7ل) =**

(أ) 11س ع + 4ص (ب)6س+3ص+5ع (ج) 6س+3ص+5ع-7ل (د)9س ص – 2ص ل

13) **إذا كان س، ص أعدادا حقيقية غير سالبة، فإنه يمكن كتابة $\sqrt{س×ص}$ على الصورة:**

(أ) س×ص (ب) $\sqrt{س}$ ×ص (ج) س×$\sqrt{ص}$ (د) $\sqrt{س}$ × $\sqrt{ص}$

14) **-3×(4-5أ) =**

(أ) -12-15أ (ب) -12+15أ^2 (ج) 12+15أ (د) -12+15أ

15) **(3س+2)(5س+3ص) =**

(أ)15س^2+9س ص+10س+6ص (ب)15س^2+6ص
(جـ)15س^2+10س+6ص (د)15س+9س ص+10س

16) صندوق مكعب الشكل طول حرفه (س^2+5)، فإن حجمه يساوي:

(أ) (س^2+5)^2 (ب) (س^2+5)^3 (ج) 2(س^2+5) (د) 3(س^2+5)

17) **ع.م.أ للحدود الجبرية : 3س ص^2 ، 15س^{2}ص ، 27س^{2}ص^3 هو:**

(أ) 3س ص (ب) 15س ص (ج) 3س^{2}ص (د) 3س ص^2

18) **يُكتب تحليل المقدار (15س ص – 4ل م +6ل ص –10س م)على الصورة :**

(أ) (3ص+2ل)(5س+2م) (ب) (3ص –2م)(5س+2ل)

(جـ) (2م –2ص)(5س+2ل) (د) (5س – 2ل)(3ص+2م)

19) **في الزوج المرتب (س،ص)، فإن ص تمثل:**

(أ) المجال (ب) المدى (ج) قاعدة الاقتران (د) بيان الاقتران

20) **عند تمثيل الاقتران ق(س)= 5س – 3 فإن الشكل الناتج يمثل :**

(أ) خطاً مستقيماً (ب) نقطة

(جـ) خط مستقيم يوازي محور السينات (د)خط مستقيم يوازي محور الصادات

21) **منحنى الاقتران ق(س)= 6س+9 يقطع محور الصادات عند النقطة:**

(أ) (0،6) (ب) (0،15) (ج) (0،9) (د) (0، 3)

22) **مدى الاقتران ق(س)={(-3،-1)،(-1،2)،(0،2)،(1،0)} هو:**

(أ) {-3،-1،0،2} (ب) {-0،2،1} ج) {-3،-0،2} د) {-2،1،1}

23) **إذا كان ق(س) اقتران خطي ميله = 3،ومقطع منحناه من محورالصادات = ـ3، فإن قاعدة الإقتران ق(س)هي :**

(أ) 3س -3 ب) -3س+3 (ج) س+3 (د) -3س ـ3

24) **(12ص+4س^2-3ل) – (7س^2+5ص) يساوي:**

(أ) 13ص س^{2}ل (ب)7ص- 3س^2+3ل (ج)5ص+9س-3ل (د) 7ص-3س^2-3ل

25) **$\sqrt{18}$ + $\sqrt{2}$ يساوي :**

(أ) $\sqrt{20}$ (ب) 20 (جـ) 4$\sqrt{2}$ (د) 6

انتهت الأسئلة

ملحق (15)

اختبار قصير (25 فقرة) من نوع الصواب والخطأ

اختبار تحصيلي في الرياضيات للصف الثامن الأساسي (الفصل الأول)

اسم الطالب...................... الشعبة.... اسم المدرسة:...........مدة الاختبار: 30 دقيقية

ضع اشارة (√) أمام العبارة الصحيحة واشارة (×) امام العبارة الخاطئة فيما يأتي:

الرقم	السـ ـــــــؤال	الإجابة
1	كل عدد نسبي هو عدد حقيقي	
2	قيمة المقدار $\sqrt{0,16} = 0,4$	
3	قيمة المقدار $\sqrt[3]{\frac{27}{64}} = \frac{3}{8}$	
4	قيمة المقدار $\sqrt[3]{125\times64} = 20$	
5	عند ضرب طرفي المتباينة $\sqrt{3} < 2$ بالعدد (-1) تصبح $-\sqrt{3} < -2$	
6	تسمى الفترة [2 ، ∞) فترة غير محدودة.	
7	طول الفترة [-2 ، 3 [يساوي (1).	
8	مجموعة حل المتباينة : 2س + 7 < 9- 4س هو: ($\frac{1}{3}$ ، ∞)	
9	قيمة المقدار $\sqrt[3]{15\frac{5}{8}}$ تساوي $\frac{2}{5}$	
10	يكتب العدد $\frac{1}{\sqrt{10}+3}$ بحيث لا يظهر الجذر في المقام على الصورة $\sqrt{10} - 3$	
11	مستطيل طوله $\sqrt{360}$ سم، وعرضه $\sqrt{40}$ سم، فإن مساحته بالسنتيمتر المربع بأبسط صورة تساوي (120)	
12	(6س + 3ص) + (5ع – 7ل) = 11س ع + 4ص	
13	إذا كان س، ص أعدادا حقيقية غير سالبة، فإنه يمكن كتابة $\sqrt{س\times ص}$ على الصورة $\sqrt{س}\times\sqrt{ص}$	
14	-3×(4-5أ) = -12-15أ	
15	(3س+2)(5س+3ص) = 15س+9س ص+10س	
16	صندوق مكعب الشكل طول حرفه (س^2+5)، فإن حجمه يساوي (س^2+5)^3	
17	ع.م.أ للحدود الجبرية : 3س ص^2 ، 15س^{2}ص ، 27س^{2}ص^3 هو 3س ص.	
18	يُكتب تحليل المقدار (15س ص – 4ل م +6ل ص –10س م)على الصورة (3ص+2ل)(5س+2م)	
19	في الزوج المرتب (س،ص)، فإن ص تمثل المدى.	
20	عند تمثيل الاقتران ق(س)= 5س – 3 فإن الشكل الناتج يمثل خط مستقيم يوازي محورالسينات.	
21	منحنى الاقتران ق(س)= 6س+9 يقطع محور الصادات عند النقطة (0،9).	
22	مدى الاقتران ق(س)={(-3،-1)،(-1،2)،(0،2)،(1،0)} هو: {-3،-1،0،2}.	
23	إذا كان ق(س) اقتران خطي ميله = 3،ومقطع منحناه من محورالصادات = –3، فإن قاعدة الإقتران ق(س) = –3س+3.	
24	(12ص+4س^2-3ل) – (7س^2+5ص) يساوي: 7ص-3س^2-3ل	
25	$\sqrt{18} + \sqrt{2}$ يساوي : $\sqrt{20}$	

ملحق (16)
مفتاح الإجابة للإختبارات الأربعة

الفقرة	اختيار من متعدد 50 فقرة				اختيار من متعدد 25 فقرة				صواب وخطأ 50فقرة	صواب وخطأ 25فقرة	الفقرة	اختيار من متعدد 50 فقرة				صواب وخطأ 50فقرة
	أ	ب	جـ	د	أ	ب	جـ	د				أ	ب	جـ	د	
1			×			×			صواب	صواب	26		×			خطأ
2		×				×			صواب	صواب	27			×		صواب
3			×					×	خطأ	خطأ	28	×				صواب
4				×	×				صواب	صواب	29		×			خطأ
5		×					×		خطأ	خطأ	30				×	صواب
6	×						×		خطأ	صواب	31				×	صواب
7			×					×	خطأ	خطأ	32	×				صواب
8		×					×		صواب	صواب	33			×		صواب
9				×			×		صواب	خطأ	34		×			خطأ
10		×					×		خطأ	صواب	35	×				صواب
11			×		×				صواب	صواب	36				×	خطأ
12	×						×		خطأ	خطأ	37			×		خطأ
13	×							×	صواب	صواب	38			×		صواب
14			×					×	خطأ	خطأ	39	×				خطأ
15				×	×				صواب	خطأ	40			×		صواب
16	×					×			خطأ	صواب	41	×				صواب
17				×	×				خطأ	صواب	42				×	خطأ
18			×			×			خطأ	خطأ	43		×			خطأ
19				×		×			صواب	صواب	44	×				صواب
20	×				×				خطأ	خطأ	45				×	صواب
21	×						×		صواب	صواب	46		×			خطأ
22				×		×			صواب	خطأ	47				×	خطأ
23		×			×				خطأ	خطأ	48			×		خطأ
24				×				×	خطأ	صواب	49			×		صواب
25	×						×		صواب	خطأ	50		×			خطأ

ملحق(17)
نتائج مطابقة فقرات الاختبارات الاربعة للنموذج اللوجستي ثنائي المعلمة لعينة من (250) مفحوص

العدد	50 فقرة				25 فقرة					50 فقرة			
النوع	اختيار من متعدد		صواب وخطأ		اختيار من متعدد		صواب وخطأ			اختيار من متعدد		صواب وخطأ	
الفقرة	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية	الفقرة	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية
1	4.6	.46	1.1	1.0	8.8	.45	10.6	.16	**26**	3.9	.69	4.7	.85
2	13.1	.07	7.8	.55	12.6	.08	4.4	.82	**27**	7.7	.17	3.2	.95
3	3.1	.37	3.1	.92	9.4	.15	2.5	.87	**28**	3.1	.79	2.8	.94
4	5.7	.45	6.2	.72	4.5	.35	9.6	.09	**29**	3.3	.86	15.8	.07
5	2.1	.90	18	.051	5.4	.61	23.1	.002	**30**	2.6	.85	5.5	.48
6	3.3	.76	5.8	.67	9.2	.32	12.8	.12	**31**	2.0	.84	10.6	.22
7	3.9	.56	5.2	.63	5.9	.65	16.4	.012	**32**	4.2	.75	6.3	.71
8	9.2	.10	26.8	.002	3.6	.73	12.8	.051	**33**	4.1	.77	6.3	.61
9	4.8	.18	4.6	.46	1.1	.95	7.3	.30	**34**	8.4	.07	23.1	.006
10	7.2	.20	9.8	.27	5.1	.65	3.6	.83	**35**	7.8	.10	4.9	.76
11	1.9	.76	5.8	.75	20.1	.01	9.6	.29	**36**	5.1	.53	9.9	.35
12	4.0	.40	10.6	.22	1.5	.90	7.7	.36	**37**	5.9	.54	7.7	.56
13	7.3	.19	1.4	.99	4.1	.66	3.5	.84	**38**	10.2	.17	4.6	.86
14	5.6	.13	12.6	.18	5.5	.60	1.3	.99	**39**	7.4	.38	10.1	.33
15	6.9	.43	5.6	.69	4.8	.57	11.2	.13	**40**	2.6	.76	16.4	.06
16	1.5	.96	15.7	.052	5.9	.66	9.6	.14	**41**	4.2	.75	5.7	.22
17	5.1	.40	15.6	.07	6.2	.51	1.7	.88	**42**	9.3	.06	8.3	.13
18	9.0	.17	2.5	.98	8.9	.26	8.3	.22	**43**	3.7	.72	12.1	.09
19	9.8	.08	7.0	.53	5.3	.62	11.3	.18	**44**	4.9	.42	4.0	.41
20	2.3	.81	10.2	.17	8.9	.34	5.7	.46	**45**	1.6	.90	3.9	.56
21	3.6	.60	14.5	.10	3.7	.71	24.3	.001	**46**	7.7	.36	0.8	.93
22	9.8	.08	7.9	.54	2.9	.89	7.9	.34	**47**	3.1	.79	4.5	.48
23	10.3	.11	5.2	.73	7.0	.22	11.5	.12	**48**	2.0	.85	2.4	.87
24	1.3	.97	4.2	.76	14.4	.07	4.1	.76	**49**	8.8	.18	11.3	.052
25	4.4	.61	4.9	.76	4.4	.73	6.9	.44	**50**	3.7	.59	11.7	.11

ملحق(18)
نتائج مطابقة فقرات الاختبارات الاربعة للنموذج اللوجستي ثنائي المعلمة لعينة من (500) مفحوص

العدد	50 فقرة				25 فقرة					50 فقرة			
النوع	اختيار من متعدد		صواب وخطأ		اختيار من متعدد		صواب وخطأ			اختيار من متعدد		صواب وخطأ	
الفقرة	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية	الفقرة	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية
1	10	.19	16.8	.053	9.7	.20	2.8	.43	**26**	14.6	.10	10.1	.18
2	11.2	.26	11.3	.25	2.8	.72	4.9	.30	**27**	6.2	.72	6.1	.63
3	5.4	.71	8.4	.40	16.1	.007	30.3	0.0	**28**	8	.53	5.9	.55
4	12.6	.18	7.6	.78	9.6	.051	20.6	.001	**29**	3.6	.94	9.9	.27
5	6.8	.65	4.3	.74	19.2	.051	9.1	.10	**30**	5.6	.58	1.8	.88
6	15.2	.055	4.9	.67	6.6	.36	1.4	.83	**31**	5.6	.78	12.2	.057
7	8.7	.46	8.6	.37	3.8	.80	3.2	.67	**32**	2.8	.97	12.2	.20
8	8.5	.48	7.4	.49	9.5	.14	2.8	.59	**33**	5.5	.59	7.3	.60
9	4.9	.67	6.5	.25	8.3	.22	9.8	.02	**34**	10.2	.33	6.5	.69
10	6	.64	3.4	.85	5.3	.378	2.5	.77	**35**	5.9	.74	9.1	.24
11	5.7	.77	7.7	.36	28.5	.002	9	.06	**36**	3.9	.91	3.9	.86
12	15.1	.09	10.5	.23	3.2	.77	7.2	.13	**37**	3.6	.93	6.9	.65
13	2.2	.98	3.6	.82	2.8	.73	5.7	.13	**38**	4.2	.89	4.9	.77
14	4.4	.88	7.8	.55	3.2	.78	.90	.92	**39**	3.6	.82	6.9	.55
15	4	.91	5.1	.75	8.2	.08	9.9	.052	**40**	11.6	.23	12.4	.19
16	7.9	.44	15.1	.058	9.2	.23	2.8	.73	**41**	9.4	.40	10.7	.21
17	12.3	.19	5.5	.79	8.3	.14	5.5	.24	**42**	7.6	.57	5.2	.64
18	6.4	.60	2.4	.96	11.3	.12	8.4	.07	**43**	3.9	.91	10.7	.29
19	2.2	.97	7.5	.58	3.3	.76	6	.20	**44**	3.2	.86	3.9	.68
20	4.2	.90	3.8	.92	5.2	.63	1.6	.85	**45**	9.7	.37	8.9	.44
21	6.8	.55	3.1	.92	11.5	.07	3.8	.28	**46**	9.6	.29	3.1	.79
22	2.4	.98	5.3	.72	1.3	.97	7.3	.12	**47**	10.7	.29	7	.54
23	3.7	.88	4	.67	4.8	.30	3.4	.49	**48**	4.6	.86	9.7	.28
24	16.8	.052	4.7	.70	6.3	.50	3.1	.53	**49**	11	.27	11.4	.24
25	10.1	.26	5	.65	5.5	.48	16.2	.003	**50**	10.9	.28	5.2	.52

ملحق(19)
نتائج مطابقة فقرات الاختبارات الاربعة للنموذج اللوجستي ثنائي المعلمة لعينة من (1015) مفحوص

العدد	50 فقرة				25 فقرة					50 فقرة			
النوع	اختيار من متعدد		صواب وخطأ		اختيار من متعدد		صواب وخطأ			اختيار من متعدد		صواب وخطأ	
الفقرة	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية	الفقرة	مربع كاي	الاحتمالية	مربع كاي	الاحتمالية
1	16.3	.051	17.1	.057	12.1	.20	7.9	.24	**26**	11.9	.22	6.6	.68
2	12.5	.18	11.9	.21	12.6	.13	12.5	.051	**27**	7.9	.54	10.7	.29
3	12.7	.17	9.6	.38	28.1	.002	52.8	0.0	**28**	22.1	.009	15.6	.07
4	14.3	.11	9.7	.37	26	.001	58.8	0.0	**29**	5.2	.81	18.9	.02
5	7.6	.58	5.1	.82	17.3	.051	7.7	.25	**30**	14.0	.12	13.5	.06
6	6.0	.73	10.0	.26	10.4	.32	10.7	.09	**31**	5.2	.82	27.8	.005
7	1.7	.99	9.2	.42	7.2	.61	11.7	.06	**32**	13.6	.13	17.9	.051
8	3.3	.95	20.6	.015	20.6	.008	17	.051	**33**	6.7	.66	3.5	.94
9	7.0	.53	16.1	.013	15.1	.57	30.8	0.0	**34**	10.4	.32	13.6	.13
10	7.3	.60	8.0	.53	6.4	.59	3.9	.69	**35**	14.1	.12	12.3	.20
11	10.8	.29	15.7	.07	35.2	.001	26.8	.002	**36**	7.4	.59	7.1	.63
12	14.4	.11	12.7	.17	9.4	.40	12.7	.052	**37**	5.1	.83	11.9	.22
13	3.4	.94	12.4	.13	19.4	.007	6.8	.14	**38**	14.4	.11	15.8	.07
14	4.0	.91	11.4	.24	13.2	.15	4.6	.46	**39**	12.0	.21	6.3	.70
15	13.3	.15	14.5	.06	9.1	.33	20.9	.002	**40**	11.0	.27	4.2	.89
16	18.0	.051	18.7	.02	22.1	.01	27.3	.001	**41**	8.0	.53	16.7	.054
17	21.2	.012	8.9	.44	15.7	.057	10.7	.09	**42**	6.7	.66	5.4	.80
18	8.2	.51	6.9	.64	15.5	.058	4.9	.43	**43**	5.8	.76	4.6	.86
19	11.5	.24	8.3	.50	9.2	.42	26.2	.002	**44**	6.8	.56	9.9	.27
20	6.2	.72	1.7	.99	17.4	.054	2.5	.87	**45**	10.0	.35	17.1	.051
21	11.2	.26	7.0	.63	9.9	.27	25.6	.001	**46**	6.6	.68	3.9	.79
22	10.2	.33	8.4	.49	7.0	.53	25.7	.001	**47**	9.6	.38	6.6	.67
23	11.2	.26	9.4	.31	8.7	.27	8.0	.159	**48**	8.0	.53	23.2	.006
24	14.4	.11	15.3	.08	7.7	.56	11.8	.06	**49**	6.4	.69	6.9	.64
25	26.4	.002	13.4	.09	18.8	.072	12.7	.052	**50**	8.2	.51	4.6	.86

THE IMPACT OF SAMPLE SIZE, TEST LENGTH, AND ITEM FORMAT ON THE ACCURACY OF ESTIMATING THE ITEM PARAMETERS, APILITY, AND TEST INFORMATION FUNCTION

By

Nabeel Mousa Jobran

Supervisor

Dr. Abdelhafez Qasem Al-Shayeb, Prof.

ABSTRACT

The aim of this study was to investigate the effect of sample size, test length, and item format on the accuracy of estimating the item parameters (difficulty, discrimination), the ability parameter, and the test information function. The sample consisted of (1015) male and female 8th grade students enrolling in ten schools which were randomly selected from the schools of Qasabat Amman during the first semester of the 2015/2016 scholastic year.

A mathematics multiple choice achievement test composed of (96) items was developed. The items were reviewed by a group of experienced specialists to verify the content validity, language accuracy, and the scientific content of the items. In order to verify the psychometric properties of the test and the effectiveness of its items, the test was applied to an initial sample of (163) students from outside the study sample. The test was divided into two forms; a long test consisting of (50) items and a short one of (25) items. Then a second true-false form of each test was derived from the long and short test forms; short multi-choice, long multi-choice, short true-false and long true-false. The four test forms were applied to the sample of the study. The answers of the subjects were

corrected and the data were entered into the memory of the computer for analysis. The total sample of the subjects was randomly divided by computer into three partial samples of different sizes; small, medium and large, each of which consisted of (250, 500, 1015) examinees respectively.

The two-parameter model was used in the framework of the item response theory to estimate the item difficulty and item discrimination parameters, the ability parameter and the test information function. The values of standard error of estimate were extracted and subjected to the analysis. A mixed One Between-Two Within Subjects repeated measures factorial ANOVA design was used to investigate the impact of sample size, test length and test form in the accuracy of both the item difficulty and item discrimination parameters, the ability parameter and the test information function.

The results of the study indicated that the difficulty parameter was more accurate when using the multi choice long test form in a large sample, while the estimation of the discrimination parameter was found to be more accurate when using the multiple choice long test form but in a small sample. The ability parameter estimation was found to be more accurate when using true false long test form in a small sample. The estimation of the test information function was found to be more accurate when using the true false long test form in a small sample.

Key Worlds: (sample size, test length, item format, estimation accuracy, item parameters, ability, test information function).